

INTRODUCTION

The proposed approach represents a dynamic and interdisciplinary domain situated at the convergence of computer vision and natural language processing (NLP). It represents a significant advancement in artificial intelligence, enabling machines to comprehend visual information and answer questions related to images, videos, or any visual content.

The idea behind VQA is to give intelligent machines the ability to interpret and reply to questions in natural language regarding the contents of the image in question. This integration of vision and language opens up a plethora of practical applications, ranging from robotics and autonomous systems to assistive technology development and vision impairment accessibility.

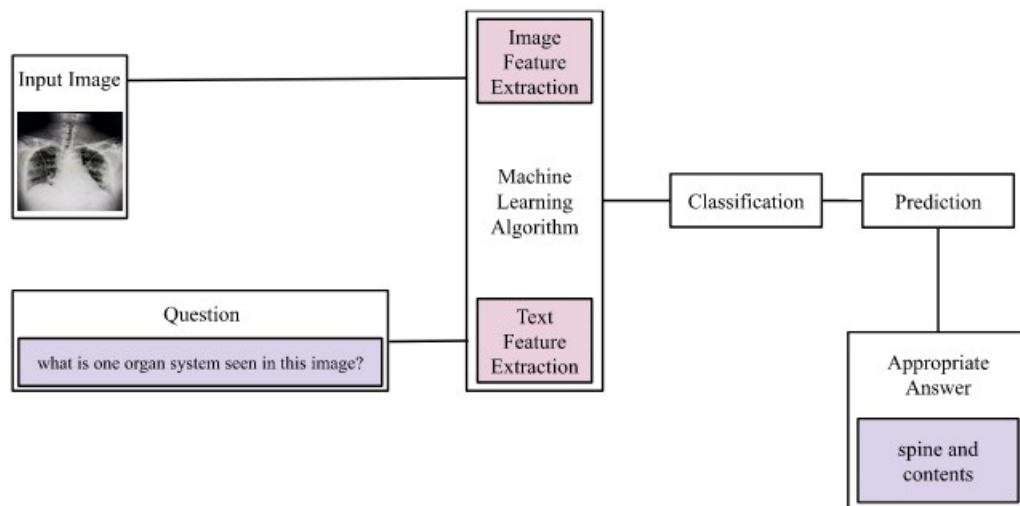


Fig. 1: Visual Question Answering system with medical image

1.1 Common discussion about VQA in healthcare domain

The primary aspiration is to develop a robust system adept at comprehending and addressing inquiries pertaining to bone characteristics and irregularities. Leveraging leading-edge for skeletal image representation algorithms, we emphasize on extracting visual features to encode the distinctive traits and patterns observed in skeletal images. Simultaneously, on the textual front, we employ advanced natural language processing algorithms to effectively process and extract meaningful insights from textual queries inputted into the Visual Question Answering (VQA) system. Recognizing that optimal communication between healthcare professionals and patients is paramount in obtaining accurate information concerning bodily conditions

and overall health, our endeavors aim to facilitate seamless interaction and understanding within this domain. Engaging in discussions concerning health matters may pose challenges for patients when faced with the technical terminology commonly employed by healthcare professionals. Successfully addressing patient concerns thus requires a multifaceted skill set, encompassing object localization, attribute identification, scene comprehension, reasoning, and counting. Attaining the requisite level of accuracy in these endeavors typically hinges on the availability of extensive labeled datasets, upon which most supervised learning algorithms heavily rely.

The fundamental challenge in VQA lies in uniting the space between the different modalities of information: visual data represented by pixels and textual data represented by natural language. To address this challenge, VQA systems employ sophisticated techniques for both visual and textual feature extraction.

1.2 Image and Textual feature representations

Image Feature Extraction: At the heart of VQA is the process of extracting significant information from healthcare skeletal images or videos. Convolutional Neural Networks (CNNs) are commonly used for visual feature extraction, allowing the system to capture intricate patterns, objects, and visual context. These visual features are then encoded into a compact representation that captures the essence of the image's content.

Textual Feature Extraction: To comprehend and process the questions, VQA systems employ natural language processing techniques. Recurrent neural networks (RNNs) or transformer-based models are frequently used to convert textual input into meaningful embeddings. These textual features encode the semantics and context of the question, enabling the system to understand the user's inquiry.

Integration of Visual and Textual Information: Once the Image and textual features are extracted, the VQA system fuses these representations to build a cohesive understanding of the input. Various fusion mechanisms, such as attention mechanisms or multimodal embeddings, are used to align the visual and textual domains effectively. This fusion facilitates the correlation between the question and the visual content, allowing the system to reason and generate appropriate answers.

Applications of Visual Question Answering: VQA systems find application in numerous real-world scenarios. For instance, they can assist autonomous vehicles in understanding their surroundings and responding to spoken instructions. In healthcare, VQA systems can aid medical professionals in diagnosing medical images or assist visually impaired individuals in understanding visual content.

Challenges and Future Directions: While significant progress has been made in VQA, challenges persist, such as handling complex questions, handling ambiguous queries, and ensuring robustness to noisy input. Researchers continue to explore novel approaches, leveraging pre-training, multimodal reasoning, and larger datasets to improve VQA system performance.

1.3 Question Answer Processing

This system incorporates interrogative response systems to assist patients, as depicted in Fig. 1. To achieve accurate output prediction, a substantial amount of data needs to be trained, including Enormous datasets containing subjective and descriptive answers to numerous questions.

However, dealing with such massive datasets poses challenges due to the computational resources required to process them effectively. Thus, a critical aspect of the system is to reduce the size of data without losing any important information. This data reduction process is crucial for optimizing machine efforts, speeding up performance, and enhancing the system's learning capabilities.

The system employs feature extraction techniques to store high-quality datasets in a new location. Utilizing pattern analysis techniques, particularly the classification algorithm, the system can classify and To accurately forecast the output answer based on the question posed about the image is to efficiently interpret and derive the most suitable response aligned with the inquiry.

During testing, the system takes an image and processes various questions to predict the appropriate answer using the data mining classification algorithm. If necessary, the system may suggest related images to further clarify the input images and questions, ensuring better support and understanding for patients using the VQA system.

Thousands of images are uploaded for training in the Visual Question Answering (VQA) system, along with corresponding relevant questions and answers for each individual image. It is common to have multiple questions and answers associated with a single image, capturing different aspects and interpretations of the visual content.

During the training process, the VQA system learns to map the images to their respective question-answer pairs, enabling it to comprehend the association between the visual features and textual information. By analyzing diverse question-answer combinations, the model acquires a more comprehensive knowledge of the image context and improves its ability to generate appropriate responses.

Having multiple questions and answers for a single image enhances the system's capability to handle variations in phrasing, question structure, and answer styles. This variability in the training data helps the VQA system generalize better to real-world scenarios where users may express their queries differently.

To achieve optimal performance, it is essential to curate a diverse and well-balanced dataset, ensuring a wide representation of questions and answers that cover various aspects of the images. This data diversity elevates the overall functionality and versatility of the VQA system in addressing a broad spectrum of visual inquiries.

In a specialized Visual Question Answering (VQA) system designed for medical images, users can pose diverse questions about the content of the images. These questions encompass a range of medical inquiries. For instance, users can ask about the identification of specific anatomical structures or abnormalities, measurements of sizes, potential diagnoses based on visual findings, recommended treatments, imaging modalities used, disease severity assessments, common medications, and treatment response evaluations.

The system aims to generate accurate and contextually relevant answers to aid healthcare professionals in medical diagnosis, treatment planning, and patient care. To achieve this, the VQA system relies on a comprehensive knowledge base of medical data, allowing it to respond with precision and provide valuable insights into the medical images. Careful validation of the answers by medical experts ensures the

appropriateness and reliability of the information provided, making the VQA system a valuable tool in the medical domain.

1.4 Different Question Types with corresponding Answers

In a Visual Question Answering (VQA) system tailored for medical images, users can ask various types of questions about the content of the images. Here are some common question types and their corresponding types of answers in a medical VQA system:

1. Anatomical identification questions:
 - Question: "What part of the body is shown in the image?"
 - Answer: A label representing the anatomical structure or body part depicted in the medical image (e.g., "liver," "lung").
2. Abnormality Detection Questions:
 - Question: "Is there any abnormality or pathology in the image?"
 - Answer: A binary response indicating the positive or negative of any medical abnormality (e.g., "Yes" or "No").
3. Quantitative Measurement Questions:
 - Question: "What is the size of the tumor in the image?"
 - Answer: A numerical value representing the size or measurement of a specific anatomical structure or pathology (e.g., "5 cm").
4. Differential Diagnosis Questions:
 - Question: "What could be the possible diagnosis based on the image?"
 - Answer: A list of potential medical conditions or diagnoses that could be associated with the visual findings in the image.
5. Treatment or Intervention Questions:
 - Question: "What is the recommended treatment for this condition?"
 - Answer: A description or a list of treatment options or medical interventions for the identified medical condition.
6. Image Modality Questions:
 - Question: "Is this image obtained from an X-ray or an MRI?"
 - Answer: A label representing the imaging modality used to acquire the medical image (e.g., "MRI").
7. Disease Severity Questions:

- Question: "How severe is the disease shown in the image?"
 - Answer: A qualitative descriptor or a numerical score indicating the severity of the identified medical condition.
8. Medication Questions:
- Question: "What medications are typically prescribed for this condition?"
 - Answer: A list of common medications used in the treatment of the identified medical condition.
9. Treatment Response Questions:
- Question: "Is the patient's condition improving after treatment?"
 - Answer: A qualitative response indicating the response of the patient's condition to a specific treatment.

It's important to note that medical VQA systems require specialized knowledge and access to accurate and reliable medical databases to provide accurate and relevant answers. The answers generated by the system should be well-vetted by medical professionals to ensure their accuracy and appropriateness for clinical decision-making. Such systems have competence to provide support to healthcare professionals in medical diagnosis, treatment planning, and patient care.

1.5 Motivation of VQA in healthcare domain

To construct a Visual QA (VQA) system in the healthcare industry is motivated by several key factors:

1. **Improved Patient Care:** A VQA system improves the healthcare outcomes and personnel make better decisions about patient care by offering easy access to essential information. For example, doctors can ask questions about medical images, such as X-rays or MRIs, to obtain insights about a patient's condition, potential diagnoses, or treatment options.
2. **Efficient Diagnosis:** Healthcare professionals often need to examine the large volumes of skeletal images to diagnose diseases or identify abnormalities. A VQA system can help streamline this process by automatically extracting relevant information from images and answering queries posed by clinicians. This can result in faster and more accurate diagnoses, ultimately improving patient outcomes.

3. **Enhanced Medical Education:** VQA systems have the potential to serve as effective educational aids for medical students, residents, and other healthcare professionals. By allowing users to ask questions about the skeletal images and receive informative responses, these systems can facilitate learning and knowledge retention in a more interactive and engaging manner.
4. **Remote Consultations:** In scenarios where specialists may not be physically present, such as in rural or underserved areas, a VQA system can enable remote consultations between healthcare providers. Clinicians can share medical images and ask questions to obtain expert opinions and guidance, leading to better patient management and care coordination.
5. **Research and Innovation:** VQA systems can support medical research efforts by providing access to huge datasets of annotated skeletal images. Researchers can use these systems to pose research questions, analyze image data, and gain insights into various medical conditions and treatments. Additionally, VQA systems can foster innovation by enabling the development of advanced image analysis algorithms and machine learning models.

Overall, the development of VQA systems in the healthcare industry is driven by the goal of improving patient care, enhancing medical education, facilitating remote consultations, and advancing medical research and innovation. By leveraging the power of computer vision and natural language processing technologies, these systems have the potential to transform various aspects of healthcare delivery and contribute to better health outcomes for patients worldwide.

1.6 Scope of VQA system in the healthcare domain

The scope of developing a Visual QA (VQA) system in the healthcare domain is vast and encompasses numerous opportunities for innovation and improvement in patient care, medical education, research, and more. Below are some key aspects that illustrate the scope of VQA systems in healthcare:

1. Diagnostic Support:

- VQA systems can assist healthcare professionals in diagnosing medical conditions by providing additional context and insights from medical images.

- Clinicians can ask questions about specific features or anomalies in images, and the VQA system can generate responses based on its understanding of the images.
- 2. Treatment Planning and Decision Support:**
- VQA systems can aid in treatment planning by providing relevant information about treatment options, potential side effects, and patient outcomes based on similar cases.
 - Healthcare providers can ask questions about treatment protocols, drug interactions, and surgical procedures, and the VQA system can provide evidence-based recommendations.
- 3. Medical Education and Training:**
- VQA systems can serve as valuable educational assets for healthcare professionals in training, residents, and other healthcare professionals by providing interactive learning experiences.
 - Learners can ask questions about skeletal images and receive detailed explanations, helping them understand complex medical concepts and procedures.
- 4. Remote Consultations and Telemedicine:**
- VQA systems can facilitate remote consultations between healthcare providers and patients, remarkably in underserved or remote areas where access to specialists may be constrained.
 - Patients can share medical images with their healthcare providers and ask questions about their condition, treatment options, and follow-up care.
- 5. Research and Data Analysis:**
- VQA systems can support medical research efforts by providing access to huge datasets of annotated skeletal images and associated clinical data.
 - Researchers can use VQA systems to analyze image data, extract meaningful insights, and identify patterns or correlations that may inform new research directions or treatment strategies.
- 6. Patient Engagement and Empowerment:**

- VQA systems can empower patients by providing them with a better understanding of their medical conditions and treatment plans.
- Patients can ask questions about their medical images, laboratory results, or treatment options, and the VQA system can provide clear and understandable explanations, helping patients make informed decisions about their health.

7. Quality Improvement and Clinical Decision Support:

- VQA systems can contribute to quality improvement initiatives by assisting healthcare providers in making more accurate and timely clinical decisions.
- By analyzing large volumes of medical image data and providing relevant information in real time, VQA systems can help clinicians optimize their workflow and improve patient outcomes.

The scope of developing VQA systems in healthcare is broad and multifaceted, encompassing various aspects of patient care, medical education, research, and quality improvement. By leveraging advances in computer vision, natural language processing, and machine learning, VQA systems have the potential to revolutionize healthcare delivery and contribute to better health outcomes for patients worldwide.

1.7 Hypothesis of VQA system in healthcare industry

The hypothesis of a Visual QA (VQA) system in the healthcare industry posits that integrating computer vision and natural language processing techniques can significantly enhance medical image analysis, clinical decision-making, patient engagement, and healthcare outcomes.

- 1. Improved Diagnostic Accuracy:** The VQA system can assist healthcare providers in accurately interpreting medical images by answering specific questions related to visual patterns, abnormalities, and diagnostic criteria. By leveraging both visual and textual information, the system aims to enhance diagnostic accuracy and reduce errors in image interpretation.
- 2. Enhanced Clinical Decision Support:** Through intelligent analysis of skeletal images and contextual understanding of clinical questions, the VQA system can provide timely and personalized decision support to healthcare

professionals. This support includes treatment recommendations, differential diagnoses, and prognostic insights based on image features and patient data.

3. **Efficient Workflow Integration:** By seamlessly integrating into existing clinical workflows and electronic health record systems, the VQA system aims to streamline image interpretation, consultation, and reporting processes. This integration facilitates efficient communication between healthcare providers, reduces turnaround time, and improves overall workflow efficiency.
4. **Empowered Patient Engagement:** The VQA system empowers patients to take an active role in their healthcare journey by providing understandable explanations and insights into their medical images. Patients can ask questions, seek clarification, and make informed decisions about their treatment plans, leading to increased engagement, satisfaction, and adherence to therapy.
5. **Facilitated Medical Education and Training:** Medical students, residents, and healthcare professionals can leverage the VQA system as a valuable educational tool for learning about medical imaging interpretation, anatomy, pathology, and clinical reasoning. The system provides interactive learning experiences, case-based tutorials, and real-time feedback to support continuous professional development.
6. **Accelerated Research and Innovation:** Researchers and scientists can leverage the VQA system to analyze large-scale medical image datasets, identify novel biomarkers, and discover patterns indicative of disease progression, treatment response, and therapeutic efficacy. This accelerated research process enables the development of innovative diagnostic tools, predictive models, and precision medicine approaches.
7. **Enhanced Quality of Care and Patient Outcomes:** Ultimately, the hypothesis suggests that the adoption of VQA systems in healthcare can lead to improvements in diagnostic accuracy, clinical decision-making, patient engagement, workflow efficiency, medical education, research productivity, and, most importantly, patient outcomes. By harnessing the power of artificial intelligence and human expertise, VQA systems have the ability to completely transform healthcare delivery and enhance patient care across various medical specialties and settings.

The hypothesis proposes that VQA systems hold immense promise for revolutionizing healthcare by leveraging advanced technologies to augment human intelligence, improve diagnostic capabilities, and ultimately enhance the quality of care provided to patients.

Here are some examples illustrating the potential applications and benefits of a Visual QA (VQA) system in the healthcare industry.

- Diagnostic Assistance
- Surgical Planning
- Patient Consultations
- Medical Education
- Remote Monitoring
- Research and Development
- Clinical Decision Support

A radiologist can use a VQA system to examine the skeletal images such as X-rays, MRIs, or CT scans. By asking specific Queries concerning the presence of abnormalities, tumor characteristics, or organ functionality, the system can provide relevant insights and assist in making accurate diagnoses. Surgeons can utilize a VQA system to better understand the anatomical structures visible in pre-operative imaging studies. By asking questions about optimal incision sites, critical landmarks, or potential complications, the system can help plan surgical approaches and anticipate challenges during procedures. During patient consultations, physicians can employ a VQA system to explain medical images to patients in a more understandable manner. Patients can ask questions about their condition, treatment options, or expected outcomes, and the system can provide personalized explanations and visual aids to facilitate shared decision-making. Medical students and residents can use a VQA system as a learning tool to enhance their understanding of medical imaging and pathology. By asking questions about image interpretation, disease mechanisms, or treatment strategies, learners can receive immediate feedback and guidance to reinforce their knowledge and skills. In telemedicine settings, healthcare providers can leverage a VQA system to remotely assess patients' conditions based on uploaded images or video consultations. By asking questions about symptom severity, treatment

adherence, or recovery progress, the system can help monitor patients' health status and provide timely interventions. Researchers can employ a VQA system to analyze large-scale medical image datasets and identify patterns associated with disease progression or treatment response. By asking questions about imaging biomarkers, genetic correlations, or clinical outcomes, researchers can gain valuable insights to inform drug development, clinical trials, and precision medicine initiatives. Clinicians can integrate a VQA system into clinical decision support systems to assist in interpreting complex skeletal images and laboratory results. By asking questions about differential diagnoses, prognostic factors, or treatment guidelines, the system can provide evidence-based recommendations to guide patient care and improve outcomes. These examples highlight how a VQA system can be applied across various healthcare scenarios to enhance diagnostic accuracy, improve patient communication, support medical education, facilitate research endeavors, and ultimately, optimize the delivery of healthcare services.

1.8 Research contribution to the society

Research contributions to society for Visual QA (VQA) systems in the healthcare domain are significant and multifaceted. Some key contributions include:

- 1. Improved Diagnostic Accuracy:** VQA systems assist healthcare professionals in interpreting medical images more accurately by providing contextual information and answering specific questions about abnormalities, anatomical structures, or disease characteristics. This can lead to more exact diagnosis and treatment strategies, hence improving patient outcomes and lowering medical errors.
- 2. Enhanced Patient Care and Communication:** VQA systems facilitate better communication between healthcare providers and patients by translating complex medical information into understandable visual representations. Patients can ask questions about their condition, treatment options, or test results, and receive personalized explanations from the system, leading to increased patient satisfaction and engagement in their care.
- 3. Efficient Medical Education:** VQA systems serve as valuable Educational assets for healthcare professionals in training, residents, and practicing clinicians by offering interactive learning experiences and real-time feedback

on medical image interpretation and diagnostic reasoning. This contributes to the continuous professional development of healthcare professionals and ensures the delivery of high-quality care.

4. **Accelerated Research and Innovation:** VQA systems enable researchers to analyze large-scale medical image datasets more efficiently and extract meaningful insights into disease mechanisms, treatment responses, and prognostic factors. By automating the process of image analysis and interpretation, these systems expedite the discovery of new biomarkers, therapeutic targets, and diagnostic algorithms, leading to advancements in medical science and technology.
5. **Accessible Healthcare Services:** VQA systems have the potential to democratize access to healthcare services by extending the reach of medical expertise to underserved populations and remote areas. Telemedicine platforms equipped with Visual QA (VQA) capabilities enable patients to receive timely consultations and diagnostic assessments from specialists without requiring physical visits, thereby reducing healthcare disparities and fostering health equity.
6. **Data-driven Decision Support:** VQA systems generate valuable insights from medical imaging data that can inform clinical decision-making and guide evidence-based practice. By integrating VQA capabilities into clinical decision support systems, healthcare providers can access relevant information and recommendations at the point of care, leading to more informed treatment decisions and better patient outcomes.

Overall, research contributions in the development and implementation of VQA systems in the medical domain have the potential to revolutionize healthcare delivery, improve patient care, and advance healthcare knowledge, ultimately benefiting society as a whole.

1.9 Application of VQA

Visual Question Answering (VQA) has numerous applications across various domains, including but not limited to:

1. Healthcare:

- **Medical Image Analysis:** VQA can assist healthcare professionals in interpreting skeletal images such as X-rays, MRIs, and CT scans by answering questions about the content of the skeletal images, aiding in diagnosis and treatment planning.
- **Clinical Decision Support:** VQA systems can provide real-time assistance to healthcare providers by answering questions related to patient data, treatment protocols, and medical literature, helping them make informed decisions.
- **Patient Education:** VQA applications can be used to create interactive educational materials for patients, allowing them to ask questions about their medical conditions, treatment options, and lifestyle changes.

2. Education:

- **Interactive Learning:** VQA can enhance traditional educational materials by allowing students to ask questions about visual content such as diagrams, charts, and graphs, facilitating deeper understanding and engagement.
- **Assessment and Feedback:** VQA systems can be used to automatically generate questions based on educational content and provide immediate feedback to students, allowing for personalized learning experiences.
- **Language Learning:** VQA applications can help language learners improve their skills by providing visual prompts and answering questions about vocabulary, grammar, and cultural context.

3. E-commerce:

- **Product Recommendation:** VQA can be integrated into e-commerce platforms to assist shoppers in finding products that meet their specific

needs and preferences by answering questions about product features, compatibility, and usage.

- **Customer Support:** VQA systems can provide automated customer support by answering questions about product specifications, pricing, shipping, and returns, improving user experience and reducing the workload of customer service representatives.
- **Visual Search:** VQA can enable visual search functionality, allowing users to find products by asking questions about their appearance, brand, and category, complementing traditional text-based search methods.

4. **Navigation and Assistive Technologies:**

- **Smart Assistants:** VQA-powered smart assistants can assist users in navigating their surroundings, providing information about landmarks, directions, and points of interest based on visual cues captured by cameras or sensors.
- **Accessibility:** VQA applications can enhance accessibility for individuals with disabilities by providing spoken answers to questions about their environment, enabling them to interact with digital and physical spaces more independently.

5. **Social Media and Content Creation:**

- **Content Moderation:** VQA systems can assist social media platforms in moderating content by automatically detecting and answering questions about potentially inappropriate or harmful visual content.
- **Content Generation:** VQA can be used to generate captions, descriptions, and tags for images and videos uploaded to social media platforms, enhancing content discoverability and engagement.

These are just a few examples of the diverse applications of Visual Question Answering across different domains. As technology advances, we should expect to see more inventive applications of VQA in the future.

1.10 **Organizations of Thesis**

This thesis report is critical for presenting your study findings in an understandable, logical, and complete manner. The following is a potential thesis organization

structure. This was followed by a literature review, which included a review of relevant material and past research in the subject of visual question answering systems in general, as well as specific applications. The application focuses on medical care and radiology imaging, with several question and answer pairings. Chapter 2 covered four types of literature reviews: a review of the visual question answer system, a review of the radiology image dataset, research on current methodology techniques and algorithms, and finally a discussion on the strategy of extracting the features of both visual and textual datasets. Then identify the issues and gaps in the following chapters. Then, discuss the problem statement and the proposed approach's objectives. Then we'll talk about key concepts, theories, and methodology. The methodology chapter provides an overview of the research design and technique used. Explanation of data collection methods, tools, and procedures. Discuss the data analysis methodologies and statistical methods used. Presentation of study results, data, and analysis. To highlight crucial findings, use tables, figures, and graphs. Discuss any surprising findings or abnormalities. Then Discussion of how to interpret results in light of the research topic or hypothesis. Results were compared to prior literature. Analyze the study's merits, shortcomings, and implications. Finally, consider the conclusion and future scope. Summary of key discoveries and their significance. Restate the research question and aims. Recommendations for future research or applications.

1.11 Problem Statement for Visual Question Answering System

In the contemporary era, various technological applications have been developed in the healthcare domain, and answering questions related to medical images is a significant concept in the medical sector. Several technologies assist in understanding the human health status through radiographic images. In many hospitals, a plethora of scanning techniques and instruments are employed to gain insights into human health. Radiology visuals consist of scanned images, and various levels of images can be obtained based on the patient's health condition. Existing approaches for medical imaging are often inadequately supported, and the collection of medical datasets poses challenges due to their limited local availability.

Present VQA models employ CNNs to extract localized feature vectors for specific regions and LSTMs to encode feature vectors for the corresponding questions.

Nevertheless, when the response encompasses two neighboring local regions in the skeletal image and the query comprises a complex sentence, the accuracy of the attention mechanisms' answers is not entirely satisfactory. The current computer-aided diagnosis technology is generally limited to a single condition. The complexity of supplementary diagnostic technology, which utilizes analysis of a singular skeletal imaging type to deliver a thorough, specific portrayal of a patient's condition akin to a clinician's diagnosis, presents a significant constraint. The existing approach is unsuitable for dealing with such complex data, and vector construction produces inferior outcomes. The current method can only remember the preceding contextual details of the question and does not have the capability to make use of subsequent information. This limitation leads to errors in extracting the question feature.

1.12 Objectives for VQA's proposed approach

1. To propose an optimal feature extraction method for radiology images in the VQA system:

- This objective aims to identify and develop a feature extraction method specifically tailored for radiology images in the visual or imaginary QA (VQA) system.
- The model should effectively capture the relevant visual information from radiology images, such as anatomical structures, abnormalities, and other diagnostic features.
- Various techniques for feature extraction, including conventional techniques and deep neural network methodologies, will be explored and evaluated.
- The optimal feature extraction method should balance computational efficiency with the ability to represent complex visual patterns present in radiology images accurately.
- Performance metrics such as feature distinctiveness, discriminability, and computational cost will be considered in selecting the optimal method.

2. To offer an automated solution for answering the user's questions in radiology images:

- This objective involves designing and implementing an automated system that can generate accurate answers to questions asked by users based on medical images.
- The system will integrate the proposed optimal feature extraction method with advanced natural language processing (NLP) techniques for question understanding and answer generation.
- It will utilize the models based on deep learning, like Recurrent Neural Networks (RNNs) or Transformers to efficiently handle textual queries and generate suitable responses.
- The system will be trained on a high dimensional dataset of skeletal images paired with corresponding question-answer pairs to learn the associations between visual features and textual queries.
- Emphasis will be placed on the machine's accuracy, speed, and growth potential to manage a broad spectrum of medical images and questions effectively.

3. To suggest an innovative method for dictionary formation concerning images and Question-Answer (QA) pairs:

- This objective aims to develop an innovative approach for developing a comprehensive dictionary that maps visual features extracted from medical images to corresponding question-answer pairs.
- The algorithm will utilize advanced machine learning and data mining techniques to automatically identify and categorize relevant visual and textual features present in the image-QA pairs.
- It will consider semantic similarities, contextual information, and relevance scores to establish meaningful associations between images and their corresponding questions and answers.
- The dictionary creation algorithm will be designed to be scalable and adaptable, allowing it to accommodate new image-QA pairs and update existing mappings over time.
- The resulting dictionary will serve as a valuable resource for the automated VQA system, facilitating accurate and contextually relevant question answers.

4. To analyze the performance of the proposed approach in comparison to established techniques and showcase the originality of the proposed strategy:

- This objective involves conducting a comprehensive evaluation of the proposed methodology against current techniques and cutting-edge methods in the field of healthcare VQA.
- Evaluation criteria such as accuracy, precision, recall, F1-score, and computational efficiency will be used to assess the effectiveness and robustness of the proposed methodology.
- Comparative experiments will be conducted using benchmark datasets and real-world medical image-QA pairs to emphasize the strengths of the proposed method.
- Statistical analysis and qualitative assessments will be performed to highlight the strengths, limitations, and unique capabilities of the proposed methodology compared with existing approaches.
- The novelty of the proposed technique will be validated through empirical results, innovative design choices, and contributions to advancing cutting-edge methods in medical VQA research.

By achieving these objectives, the proposed research aims to advance the field of medical VQA by introducing novel methodologies, algorithms, and systems that enable accurate, automated, and contextually relevant question answering based on radiology images.