INDEX

CHAPTER- I INTRODUCTION		**1 – 19**
1.1	Common discussion about VQA in healthcare domain	1
1.2	Image and Textual feature representations	2
1.3	Question Answer Processing	3
1.4	Different Question Types with corresponding Answers	5
1.5	Motivation of VQA in healthcare domain	6
1.6	Scope of VQA system in the healthcare domain	7
1.7	Hypothesis of VQA system in healthcare industry	9
1.8	Research contribution to the society	12
1.9	Application of VQA	14
1.10	Organizations of Thesis	15
1.11	Problem Statement for Visual Question Answering System	16
1.12	Objectives for VQA's proposed approach	17
CHAPTER- II REVIEW OF LITERATURE		**20 - 52**
2.1	Review of relevant literature and previous research	20
	2.1.1 Review on Visual Question Answer	20
	2.1.2 Review on Radiology Image Datasets	25
	2.1.3 Research on Current Methodology with existing techniques and Algorithms	31
	2.1.4 Survey on visual and textual Feature Extraction Technique	37
2.2	Identification of challenges and gaps in existing research	46
	2.2.1 Challenges in VQA System	46
	2.2.2 Inhibitions of Datasets	47
	2.2.3 Drawbacks on various techniques	48
2.3	Gap Identification from Existing Research	49
CHAPTER-III METHODOLOGY		**53 – 102**
3.1	Description of the research design and methodology used	53
	3.1.1 Morphology of a Radiology Image	55
3.2	Explanation of data collection methods, tools, and procedures	57

3.3	Current Methodology on Both Visual and Textual Feature Extraction Techniques	58
	3.3.1 Linear Classifier	58
	3.3.2 Traditional Neural Network	59
	3.3.3 The prime Idea behind a Convolutional Neural Network for image feature extraction	60
	3.3.4 Deep Belief Network model, the feature extraction of skeletal image	74
	3.3.5 The deep insights of Region-based Convolutional Neural Network (R-CNN) for visual classification	80
	3.3.6 Faster Region-based Convolutional Neural Network (R-CNN)	84
	3.3.7 Long Short-Term Memory networks for question answering system	90
3.4	The proposed Approach for Visual Question Answering System (VQAS) for skeletal Images	93
	3.4.1 Skeletal image Feature Extraction using Block_12_add Faster R-CNN (B12- FRCNN) algorithm	94
	3.4.2 The proposed approach Kai-Bi-LSTM for Question Answering feature extraction	98
CHAPTER- IV EXPERIMENTAL RESULT AND ANALYSIS		**103 - 139**
4.1	Experimental Result on VQA	103
4.2	Code Implementation	117
4.3	Comparison analysis of Feature extraction techniques	120
4.4	Snapshot on demonstration	125
CHAPTER- V CONCLUSION AND FUTURE ENHANCEMENT		**140 – 142**
REFERENCE		**143 - 155**
PUBLICATIONS		
CERTIFICATES		
PLAGIARISM REPORT		

LIST OF TABLE

Table No.	Particulars	Page No.
1	Descriptive Statistics of Existing visual and textual dataset	28
2	Correlation of Existing visual and textual dataset	28
3	Covariance of Existing visual and textual dataset	29
4	Cumulative Frequency of Existing visual and textual dataset	29
5	Describes the publications obtained and reviewed during the current survey	41
6	Most Frequent Answers for various Question Type and Answer count	110
7	Total count of Visual and Textual dataset	112
8	The table presents the performance metrics for a certain method across different evaluation criteria	120
9	The accuracy achieved by different classification algorithms	122
10	The table shows the F-measure achieved by different algorithms	123
11	The table compares the accuracy of both Proposed and Current algorithms	124
12	The comparative analysis between Existing CNN and Existing BiLSTM with Proposed BiLSTM	125

LIST OF FIGURE

Fig. No.	Particulars	Page No.
1	Visual Question Answering system with medical image	1
2	Trends in Visual Question Answering Research	21
3	Histogram of existing collective dataset with its frequency	30
4	Total data present in various image and question answer pair dataset	30
5	Insight of Literature Survey	45
6	Types of Radiology Image	56
7	Traditional Neural Network with a Single Layer	60
8	RGB Image and Gray scale Image to find features of the image	60
9	3 × 3 pixel object	61
10	Structure of CNN process from Input object to final output object with feature extraction	61
11	The value of Black and white colors 0 to 255	61
12	How the convolution operation take place with 6 × 6 Pixel images and 3 × 3 image	62
13	Black and white input image for convolution operation to identify the vertical edge	62
14	Black and White image with 3 × 3 px vertical edge filter the output is in 4 × 4px	63
15	Output image from above figure after 3 × 3 px convolution with 6 × 6 px	63
16	Image 3 × 3 px horizontal edge filter	64
17	n × n × 1 for Gray scale image with f × f × c number of filters the output will be multiple images	64
18	Single filter convolution with single output	66
19	Multiple filter convolution with multiple output	66
20	Padding in Convolutional Neural Network	66
21	Single Filter Convolution layer using non linear function and bias with single image as output	69

Fig. No.	Particulars	Page No.
22	Multiple Filter Convolution layer using non linear function and bias with multiple output So, in short, we convert it as	70
23	Overall Convolution layer using non linear function and bias with multiple output in one frame	71
24	The Architecture of Convolutional Neural Network	73
25	The Traditional Architecture of Artificial Neural Network	74
26	Two-layer probabilistic neural network, RBM Architecture	76
27	Overall architecture of a Deep Belief Network	78
28	Faster Region-based Convolutional Neural Network (R-CNN)	84
29	Understanding the Fast RCNN and Faster RCNN Architecture	90
30	Architecture of Long Short-Term Memory (LSTM) networks	93
31	B12-FRCNN, the "block_12_add" layer	97
32	Block diagram of QA System in the Training phase	102
33	Block diagram of QA System on Testing Phase	102
34	Confusion Matrix	106
35	Most frequent answers from different question types	111
36	Different methods of question for various question types	112
37	Total count of Image and questions from CLEF Image Retrieval and Classification Task 2019 Dataset	112
38	Total number of data from both question and answer for various types	114
39	Count of boolean questions for Modality	114
40	Various types of question for Modality question answering data type	115
41	Various types of question for Plane question answering data type	115
42	Various types of question for Organ question answering	116

Fig. No.	Particulars	Page No.
	data type	
43	Various types of question for Abnormality question answering data type	116
44	Overall count of different question types	116
45	Sample code function on Existing CNN algorithm	117
46	Construction of Existing DBN structure for Medical Images	118
47	Code function on Proposed Kai_BiLSTM	120
48	The performance metrics for a certain method across different evaluation criteria	121
49	The accuracy achieved by various classification techniques for current methodology	122
50	F-measure achieved by different algorithms	123
51	Comparison with existing algorithm and proposed algorithm	124
52	The Comparison Result of Proposed and Existing algorithms	125
53	Starting page of Application	125
54	Load the Training dataset both Image and Question Answer pair	126
55	Feature Extraction for Image dataset	126
56	Pre-process the Question dataset	127
57	Pre-process for Answer Dataset	127
58	Done with Preprocessing for both Question and Answer dataset	128
59	Word Embedding for Question Dataset using BERT model	128
60	Word Embedding for Answer Dataset using LBERT model	129
61	Training the dataset for both visual and textual dataset	129
62	Load the Testing dataset both Image and Question Answer	130

Fig. No.	Particulars	Page No.
	pair	
63	Feature Extraction for Testing image dataset	130
64	Pre-processing for Question Datasets using LBERT model	131
65	Testing both visual and textual dataset	131
66	Precision measure for proposed and existing models	132
67	Recall measure for proposed and existing models	132
68	Classification of FMeasure for proposed and existing models	133
69	Classification Sensitivity measure for proposed and existing models	133
70	Classification Specificity measure for proposed and existing models	134
71	True Positive Rate (TPR) measure for proposed and existing models	134
72	Positive Predictive Value (PPV) measure for proposed and existing models	135
73	Measure False Negative Rate (FNR) for proposed and existing models	135
74	To display Graphs and Tables	136
75	Load the skeletal Image	136
76	Feature Extraction for Loaded Image using B12-FRCNN	137
77	Users Input Question	137
78	Preprocess the Input Question	138
79	Word Embedding for Input Question	138
80	Classification using Kai-BiLSTM model	139
81	Predicated answer with calculated score value for the answer	139