Table of Contents

PageNo.

- Declaration by Research Scholar–Originality of Research Work……. II
- Certificate of Supervisor…… III
- Thesis Approval Form……... IV
- Declaration by Research Scholar – Submission of Thesis……..... V
- Acknowledgement ……..... VI
- Table of Contents……... VII
- List of Figures…… XII
- List of Tables……. XIV
- Abstract…….. XV

1. Introduction to Voice Recognition System……... 1

1.1 Introduction to Speech Processing System…….. 2

1.2 Types of Speech Processing……. 3

1.2.1 Speech Recognition……..... 3

1.2.2 Voice Recognition……... 3

1.2.2.1 Voice Identification……..... 4

1.2.2.1 Voice Verification……..... 4

1.2.3 Speech Synthesis...…….. 5

1.2.4 Speech Enhancement……... 5

1.2.5 Emotion Recognition…….. 6

1.2.6 Language Recognition...……. 6

1.3 Gujarati Language and Dialects……. 6

1.4 Statement of the Problem……. 9

1.5 Need of Proposed Research…….. 10

1.6 Objectives of the Research…… 12

1.7 Scope of the Study……..... 13

1.8 Limitations of the Study…… 14

1.9 Structure of the Research……. 15

References…….. 16

2. Literature Review……..……. 19

2.1 Introduction…… 19

2.2 Previous Work in Field of Speech Recognition System…….. 19

2.3 Previous Work in Field of Voice Recognition System…….. 28
2.4 Previous Work in Field of Voice Recognition for Indian Language…… 32
2.5 Previous Work in Field of Voice Recognition for Gujarati Language…. 37
2.6 Previous Work Related to Feature Extraction Techniques……. 45
References……... 53

3. Study and Analysis of Gujarati Voice for Structural Feature Extraction and Classification……. 65

3.1 Introduction to Gujarati Language……..... 65
3.2 Gujarati Speech Features…… 65
3.2.1 Voice Biometrics…….. 65
3.2.1.1 Pitch (High vs. Low Pitch) …….. 66
3.2.1.2 Timbre…….. 67
3.2.1.3 Speech Rate (Tempo)…… 68
3.2.1.4 Rhythm (Tempo and Flow)……... 68
3.2.1.5 Emotion in Speech (Expressiveness)…… 70
3.2.2 Phonetic Features…….. 73
3.2.2.1 Vowel System……... 73
3.2.2.2 Consonant System……. 74
3.2.3 Prosodic Features…….. 76
3.2.3.1 Stress Patterns (Stress and Accent)……... 76
3.2.3.2 Intonation Patterns (Pitch and Tone)……..... 77
3.2.3.3 Pauses and Silences……... 79
3.2.4 Syntactic Features…… 80
3.2.4.1 Word Order……... 80
3.2.4.2 Postpositions…… 81
3.2.5 Linguistic Features…… 81
3.2.6 Speech Articulation……... 83
3.2.6.1 Dialects Variations…… 83
3.2.7 Voice Quality…… 85
3.2.7.1 Breathiness & Roughness……. 85
3.2.7.2 Voice Clarity & Speech Rate…… 86
3.3 Introduction to Feature Extraction Techniques……... 89
3.4 Different Feature Extraction Techniques for Gujarati Dialects……... 90

3.4.1 Mel-Frequency Cepstral Coefficients (MFCC)…….. 91
3.4.1 Linear Prediction Cepstral Coefficients (LPCC)…… 92
3.4.2 Perceptual Linear Prediction (PLP)…… 93
3.4.3 Discrete Wavelet Transform (DWT)…….. 94
3.4.4 Fast Fourier Transform (FFT)…… 94
3.4.5 Relative Spectral (RASTA)……... 95
3.5 Evaluating Best Feature Extraction Technique for Gujarati Dialects….. 96
3.6 Evaluating Different Classifier for Voice Recognition……... 98
3.6.1 Hidden Markov Model (HMM)……. 99
3.6.2 Gaussian Mixture Model (GMM)…….. 100
3.6.3 K-Nearest Neighbors (KNN)……. 101
3.6.4 Support Vector Machine (SVM) ……... 101
3.6.5 Convolution Neural Network (CNN) ……. 102
3.6.6 Long Short-Term Memory (LSTM)…… 103
3.6.7 Time Delay Neural Networks (TDNN) ……. 103
3.6.8 Recurrent Neural Networks (RNN)…… 104
References…….. 105

4. Proposed Framework & Voice Recognition Model for Vernacular Gujarati Dialects…… 107

4.1 Introduction…….. 107
4.2 Proposed Framework for Vernacular Voice Recognition of Gujarati Dialects……... 108
4.2.1 Input Layer…….. 109
4.2.2 Preprocessing Layer……... 109
4.2.3 Processing Layer……. 110
4.3 Voice Recognition Model for Gujarati Dialects…… 111
4.3.1 Speech Data Collection…… 113
4.3.2 Speech Data Organization……... 113
4.3.3 Voice Recognition Model…… 114
4.3.3.1 Data Preprocessing…….. 115
4.3.3.2 Feature Extraction……..... 117
4.3.3.3 Classification……..... 118
4.3.3.4 Display the Results……... 118

References…… 119

5. Component Development of Voice Recognition Model for Vernacular Gujarati Dialects…….. 121

5.1 Introduction……... 121

5.2 Speech Data Collection……. 121

5.2.1 Privacy and Security Measures in Voice Data Collection……. 122

5.2.2 Sources of the Data……... 123

5.2.3 Samples of Collected Voice Data (Unstructured)……. 124

5.2.4 Nature of the Data (Unstructured Data)…… 124

5.3 Data Organizations…… 125

5.3.1 Convert the audio files into standard format……. 125

5.3.2 Categorization of Unstructured Data…… 125

5.3.3 Define a Directory Structure……. 126

5.3.4 Annotations and Metadata…… 128

5.4 Designing User Interface for Voice Recognition (MVR) Model of Gujarati Dialects……. 129

5.4.1 Features of Voice Recognition (MVR) Tool…… 129

5.4.2 Component of Voice Recognition (MVR) Tool……... 131

5.4.3 Sample Outputs…… 133

5.5 Pseudo Code used to Build Voice Recognition Tool and Model for Vernacular Gujarati Directs……..... 137

References…… 148

6. Performance Analysis, Results, Conclusion and Future Scope for Extension of Research Work…….. 149

6.1 Introduction……... 149

6.2 Performance Analysis of Vernacular Voice Recognition Model for Gujarati Dialects…….. 149

6.2.1 Speaker Wise Result Analysis for Voice Recognition……. 150

6.2.2 Gender Wise Recognition Analysis……. 157

6.2.3 Dialect Wise Recognition Analysis……. 159

6.2.4 Duration Wise Accuracy Analysis…… 160

6.3 Results & Discussions……... 161

6.4 Conclusion……. 162

6.5 Future Scope…….. 162
References……... 163
❖ Appendix…….. 164

Table of Figures

Figure No.	Figure Title	Page No.
1- 1	Speech Processing Taxonomy…..……	2
1- 2	Map showing Region wise Dialect Diversity in Gujarat ……..…..	8
3- 1	Pitch Variation in Neutral Sentence ……..……..……..…….	66
3- 2	Pitch Variation in Spoken Sentence Conveying Regards …..…….....	66
3- 3	Pitch Variation in Assertive Sentence …..……..……..…….....	67
3- 4	Increasing Speech Rate with Blended Words …..……..……..……..	68
3- 5	Pitch Intonations of sentence Conveying Excitement ……..…….....	70
3- 6	Pitch Intonations of sentence Conveying Displeasure……..…….....	70
3- 7	Pitch Intonation of Sentence Conveying Sadness……..……..……..	70
3- 8	Rising Intonation Pattern of Spoken Sentence ..…..……..……..….	78
3- 9	Falling Intonation Pattern of Spoken Sentence ……..……..……...	78
3- 10	Neutral Intonation Pattern of Spoken Sentence ..……..…….....	79
3- 11	Speech Rate of Sentence having Pause …..……..…….....	80
3- 12	Linguistic Features of Gujarati for Different Gender ..…….....	81
3- 13	Pronouns in Gujarati Language ..……..……..…….....	82
3- 14	Compound Words ……..……..……..……..……..……..……..	82
3- 15	Pitch Intonation of Sentence with Breathy Voice ..……..	86
3- 16	Pitch Intonation of Sentence with Rough Voice …..……	86
3- 17	Speech Rate of Informal Sentence……..……..……..……	87
3- 18	Speech Rate of Formal Sentence ..…..……..……..….....	87
3- 19	Speech Rate of a Sentence Signalling Question ……..……..……..	88
3- 20	Fundamental Process of Feature Extraction……..……..……..……..	89
3- 21	MFCC Feature Extraction Process ……..……..……..……..……	91
3- 22	LPCC Feature Extraction Process ..…..……..……..……..……..	92
3- 23	PLP Feature Extraction Process ……..……..……..……..…….....	93
3- 24	DWT Feature Extraction Process ..…..……..……..……..…….....	94
3- 25	FFT Feature Extraction Techniques ……..……..……..……..……..	95
3- 26	RASTA Feature Extraction Process …..……..……	95

4- 1	Proposed Framework for Vernacular Voice Recognition Model of Gujarati Dialects ..……...……...……...…….	108
4- 2	Voice Recognition System for Vernacular Gujarati Dialects……	112
4- 3	Vernacular Voice Recognition Model of Gujarati Dialects (MVR Model)……	115
4- 4	Data Preprocessing Operations ..……...……...…….....	116
5- 1	Samples of collected data from various public resources ……	124
5- 2	Directory Structure of 'vernacVoiceData' Dataset …...……	127
5- 3	Screenshot of Voice Recognition (MVR) Tool ……...…….....	130
5- 4	Title Bar of MVR Tool …...……...……...…….....	131
5- 5	Navigation Bar of MVR Tool…...……...…...…….....	131
5- 6	Input Region of Recognize Voice window of MVR Tool ...…….....	131
5- 7	Output Region of Recognize Voice window of MVR Tool …….....	132
5- 8	Enrol New Voice window of MVR Tool ……...……...…….....	132
5- 9	Sample Output displaying Predicted Speaker along with gender, dialect & confidence of voice recognition of uploaded audio ……..	134
5- 10	Sample output of voice sample of speaker033 of Kutchhi dialect ……	134
5- 11	Sample output of voice sample of random speaker of any dialect ..…..	134
5- 12	Sample output of voice sample of Kathiyawadi dialect ……...…….	135
5- 13	Sample output of voice sample of Surti dialect ……...……...…….....	136
6- 1	Speaker Wise Recognition Analysis of Kathiyawadi Dialects ...……..	152
6- 2	Speaker Wise Recognition Analysis of Kutchhi Dialects ...……...…..	153
6- 3	Speaker Wise Recognition Analysis of Standard Gujarati Dialects…..	156
6- 4	Speaker Wise Recognition Analysis of Surti Dialects …...……...……	157
6- 5	Gender Wise Recognition Analysis …...……...……...……...……..	158
6- 6	Gender Wise Accuracy Analysis …...……...……...……...……	158
6- 7	Dialect Wise Recognition Analysis …...……...……...…….....	159
6- 8	Duration Wise Accuracy Analysis …...……...……...…….....	160

List of Tables

Table No.	Table Title	Page No.
3- 1	Table showing the Examples of Short Vowels	73
3- 2	Table showing example of Long Vowels	74
3- 3	Table Showing Diphthongs examples	74
3- 4	Table Showing Nasalized Vowels	74
3- 5	Table showing Example of Consonant System	75
3- 6	Prosodic Features of Gujarati Speech	76
3- 7	Region Wise Speech Accent Examples	83
3- 8	Performance Measures Obtained from Various Feature Extraction Techniques	97
5- 1	Table shows sample metadata of audio files	128
5- 2	Pseudo Code for MVR_main	137
5- 3	Pseudo Code for MVR_noise_reduction	138
5- 4	Pseudo code for MVR_silence_removal	139
5- 5	Pseudo Code for MVR_audio_segmentation	140
5- 6	Pseudo Code for MVR_audio_segmentation	141
5- 7	Pseudo Code for MVR_resample_audio_files	142
5- 8	Pseudo Code for MVR_extract_flat_mfcc	143
5- 9	Pseudo Code MVR_load_data_ml	144
5- 10	Pseudo Code MVR_train_svm_classifier	145
5- 11	Pseudo Code MVR_predict_speaker_svm	146
6- 1	Result Analysis of each Spekaer of Kathiyawadi Dialect	151
6- 2	Result Analysis of each Spekaer of Kutchhi Dialect	152
6- 3	Result Analysis of each Spekaer of Standard Gujarati Dialect	155
6- 4	Result Analysis of each Spekaer of Surti Dialect	156
6- 5	Gender Wise Recognition Analysis	158
6- 6	Dialect Wise Recognition Analysis	159
6- 7	Duration Wise Accuracy Analysis	160