

Tabla de Contenido

1. Introducción	1
1.1. Identificación y Formulación de Problema	1
1.2. Objetivos del Trabajo de Título	2
1.3. Estructura de la memoria	2
2. El problema de separación de fuentes y <i>Deep Learning</i>	3
2.1. Marco Teórico	3
2.1.1. Interacción humano-robot por voz	3
2.1.2. <i>Deep Learning</i>	4
2.1.2.1. Redes convolucionales	5
2.1.2.2. Redes recursivas	7
2.1.2.3. <i>Padding</i>	8
2.1.2.4. Concatenar	9
2.1.2.5. <i>Kernel dilatation</i>	10
2.1.2.6. <i>Skip connection</i>	10
2.1.2.7. Máscaras	11
2.1.3. <i>Source Separation</i>	11
2.1.3.1. Reverberación	12
2.1.3.2. <i>Singal to Noise Ratio</i> (SNR) y <i>Scale Invariant-SNR</i> (SI-SNR)	13
2.1.3.3. <i>Weigth Prediction Error</i> (WPE)	13
2.1.3.4. <i>Room Impulse Response</i> (RIR)	13
2.1.3.5. <i>Direction Of Arrival</i> (DOA)	14
2.1.3.6. <i>Interchannel Phase Difference</i> (IPD)	14
2.1.3.7. <i>Directional Feature</i> (DF)	14
2.1.4. <i>Beamforming</i>	15
2.1.4.1. <i>Sum and Delay</i>	16
2.1.4.2. <i>Generalized Eigenvector</i> (GEV)	17
2.1.4.3. <i>Minimum Variance Distortionless Response</i> (MVDR)	17
2.1.5. <i>Automatic Speech Recognition</i> (ASR)	18
2.1.5.1. <i>Word Error Rate</i> (WER)	19
2.2. Estado del arte	19
2.2.1. Estudios de separación de fuentes	19
2.2.2. Técnicas de separación de fuentes	23
2.3. Contribución	25
3. Separación de fuentes aplicado a la interacción humano-robot	27
3.1. Base de datos	27

3.1.1.	Base de datos con movimiento simulado	27
3.1.2.	Base de datos con movimiento real	30
3.1.2.1.	Base de datos THRI	30
3.1.2.2.	Base de datos VS	33
3.2.	Aspectos Metodológicos	35
3.2.1.	Generación de data simulada	35
3.2.2.	Pre-procesamiento de los datos	37
3.2.2.1.	Base de datos BD_{sim}	37
3.2.2.2.	Base de datos BD_{THRI} y BD_{VS}	37
3.2.3.	<i>Beamforming</i> y <i>Sum Without Delay</i> (SWD)	38
3.2.4.	Modelo de <i>source separation</i>	38
3.2.4.1.	Red convolucional TCN/CBP	38
3.2.4.2.	Red neuronal Deep-Gev	39
3.2.4.3.	Red neuronal Deep-MVDR-steering	41
3.2.5.	ASR STE-CROS	42
3.2.6.	Métricas de desempeño	43
4.	Discusión de Resultados	44
4.1.	Resultados	44
4.1.1.	Replicación resultados investigación LPTV con red TCP/CBP	44
4.1.2.	Replicación de <i>beamforming</i> con base de datos real THRI	45
4.1.3.	Replicación de <i>beamforming</i> con base de datos real VS	46
4.1.4.	Resultados de red Deep-GEV/Deep-MVDR-steering sobre base de datos real	46
4.2.	Análisis de resultados	48
5.	Conclusiones	52
	Bibliografía	54