

# Tabla de Contenido

|  |           |
|--|-----------|
| <b>1. Introduccion</b>   | <b>1</b>  |
| 1.1. Antecedentes Generales . . . . .                              | 1         |
| 1.2. Identificación y Formulación del Problema . . . . .           | 2         |
| 1.3. Objetivos del Trabajo de Título . . . . .                     | 2         |
| 1.4. Alcances del Proyecto . . . . .                               | 3         |
| 1.5. Justificación del Proyecto . . . . .                          | 3         |
| 1.5.1. Relevancia del Problema . . . . .                           | 3         |
| 1.5.2. Laboratorio de Procesamiento y Transmisión de Voz . . . . . | 4         |
| 1.5.3. Agradecimientos . . . . .                                   | 5         |
| 1.6. Estado del Arte . . . . .                                     | 5         |
| <b>2. Marco Teórico</b>  | <b>8</b>  |
| 2.1. Hidden Markov Model . . . . .                                 | 8         |
| 2.1.1. Algoritmo de Viterbi . . . . .                              | 10        |
| 2.2. Filtro Triangular de Frecuencias . . . . .                    | 12        |
| 2.3. Gaussian Mixture Model . . . . .                              | 13        |
| 2.4. Deep Neural Network . . . . .                                 | 15        |
| 2.5. Passive Acoustic Monitoring . . . . .                         | 18        |
| 2.6. Vocalizaciones de Ballenas . . . . .                          | 20        |
| 2.6.1. Otros . . . . .   | 26        |
| 2.7. Software Kaldi . . . . .                                      | 28        |
| 2.8. Espectrograma . . . . .                                       | 28        |
| 2.9. FFT . . . . .   | 29        |
| 2.10. Terminología procesamiento de la voz . . . . .               | 30        |
| 2.10.1. Utterance . . . . .  | 30        |
| 2.10.2. Evento . . . . .   | 31        |
| 2.10.3. Frame . . . . .  | 31        |
| 2.10.4. Transcripción . . . . .                                    | 31        |
| 2.10.5. Alineamiento . . . . .                                     | 31        |
| 2.11. Ventana de Hamming . . . . .                                 | 32        |
| 2.12. Autoencoder . . . . .  | 33        |
| <b>3. Datos Disponibles</b>  | <b>36</b> |
| <b>4. Metodología</b>  | <b>38</b> |
| 4.1. Preprocesamiento . . . . .                                    | 39        |
| 4.1.1. Procesamiento de las grabaciones . . . . .                  | 40        |
| 4.1.2. Procesamiento de las anotaciones . . . . .                  | 42        |

|           |   |           |
|-----------|---|-----------|
| 4.2.      | División de la base de datos . . . . .                                    | 42        |
| 4.3.      | Extracción de <i>features</i> . . . . .                                   | 43        |
| 4.3.1.    | Banco de filtros triangulares . . . . .                                   | 45        |
| 4.3.2.    | HMM . . . . .   | 46        |
| 4.3.3.    | Símbolos observables . . . . .  | 47        |
| 4.3.4.    | Espacio de estados . . . . .  | 47        |
| 4.3.5.    | Probabilidades de transición . . . . .                                    | 50        |
| 4.3.6.    | Probabilidades iniciales . . . . .  | 50        |
| 4.3.7.    | Probabilidades de observación . . . . .                                   | 51        |
| 4.4.      | Gaussian Mixture Model . . . . .  | 51        |
| 4.4.1.    | Determinación de parámetros de las gaussianas iniciales . . . . .         | 52        |
| 4.4.2.    | Ajuste de las gaussianas . . . . .  | 53        |
| 4.5.      | Deep Neural Network . . . . .   | 54        |
| 4.5.1.    | Pre-entrenamiento de la red DNN ( <i>Autoencoder</i> ) . . . . .          | 55        |
| 4.5.2.    | Estructura del modelo DNN . . . . .                                       | 56        |
| 4.5.3.    | Primera red neuronal profunda . . . . .                                   | 57        |
| 4.5.4.    | Segunda, tercera y cuarta DNN . . . . .                                   | 57        |
| 4.6.      | Método de cálculo de resultados . . . . .                                 | 58        |
| <b>5.</b> | <b>Resultados</b>   | <b>60</b> |
| 5.1.      | Base de Datos . . . . .   | 60        |
| 5.2.      | Gaussian Mixture Model . . . . .  | 61        |
| 5.3.      | Deep Neural Network . . . . .   | 64        |
| <b>6.</b> | <b>Análisis de Resultados</b>   | <b>67</b> |
| 6.1.      | Base de Datos . . . . .   | 67        |
| 6.2.      | Gaussian Mixture Model . . . . .  | 68        |
| 6.3.      | Deep Neuronal Network . . . . .   | 69        |
| 6.4.      | Comparación entre resultados con GMM y con DNN . . . . .                  | 70        |
| <b>7.</b> | <b>Conclusión</b>   | <b>72</b> |
|           | <b>Bibliografía</b>   | <b>75</b> |
|           | <b>Anexo A. Hidden Markov Model</b>                                       | <b>79</b> |
| A.1.      | Algoritmo <i>forward</i> . . . . .  | 79        |
| A.2.      | Algoritmo de Baum-Welch . . . . .   | 80        |
|           | <b>Anexo B. Filtro Triangular de Frecuencias</b>                          | <b>82</b> |
| B.1.      | Mel Frequency Ceptrum (MFC) . . . . .                                     | 82        |
|           | <b>Anexo C. Gaussian Mixture Model</b>                                    | <b>83</b> |
| C.1.      | Algoritmo de esperanza-maximización (EM) . . . . .                        | 83        |
|           | <b>Anexo D. Deep Neuronal Networks</b>                                    | <b>85</b> |
| D.1.      | <i>Backpropagation</i> y el algoritmo de gradiente descendiente . . . . . | 85        |