



UNIVERSIDAD DE VALLADOLID

ESCUELA TÉCNICA SUPERIOR DE
INGENIEROS DE TELECOMUNICACIÓN

Análisis del uso de algoritmos de *Deep Learning* en un Sistema biométrico multimodal

Trabajo Fin de Grado

Grado en Ingeniería de Tecnologías
Específicas de Telecomunicación.
Mención Sistemas de Telecomunicación

Autor:

Alberto García García

Tutora:

Lara del Val Puente

Valladolid, marzo 2024

Título	Análisis del uso de algoritmos de <i>Deep Learning</i> en un Sistema biométrico multimodal
Autor	D. Alberto García García
Tutora	Dra. Lara del Val Puente
Departamento	Teoría de la Señal y Comunicaciones

Tribunal

Presidente	Alberto Izquierdo Fuente
Secretaria	Lara del Val Puente
Vocal	Juan José Villacorta Calvo
Suplente 1	Ramón de la Rosa
Suplente 2	Alonso Alonso Alonso

Agradecimientos

Resumen

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Palabras clave

Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Keywords

Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Índice

Agradecimientos	I
Resumen	II
Abstract	III
1. Introducción	2
1.1. Estado del arte	2
1.2. Objetivos	2
1.2.1. Objetivos generales	2
1.2.2. Objetivos específicos	2
1.3. Metodología	2
2. Marco teórico	3
2.1. Arrays de micrófonos	3
2.2. Características acústicas de insectos xilófagos	3
2.3. Redes neuronales	3
2.3.1. Redes neuronales tipo transformer	3
2.4. Herramientas de desarrollo	4
2.4.1. LabVIEW	4
2.4.2. Python	4
3. Procesado de señales	5
3.1. Adquisición de señales	5
3.2. Segmentación de señales	5
3.3. Conformación de las señales	6
3.4. Creación de la base de datos (etiquetado)	6
3.5. Análisis de los datos	6
3.6. Cálculo de espectrogramas	6
4. Desarrollo de la red neuronal	7
4.1. <i>Transfer learning</i>	7
4.2. <i>Fine tuning</i>	7
4.3. Diseño de las nuevas capas	7
4.4. Entrenamiento de la red	7
5. Resultados	8
6. Conclusiones	9
Bibliografía	10
Anexos	12

Índice de figuras

2.1. Arquitectura de un Transformer [6].	4
3.1. Número total de detecciones por captura con diferentes valores de guarda. .	5

Índice de tablas

Capítulo 1

Introducción

1.1. Estado del arte

Artículos que conforman el estado del arte: [\[1, 2, 3, 4\]](#)

1.2. Objetivos

1.2.1. Objetivos generales

1.2.2. Objetivos específicos

1.3. Metodología

Capítulo 2

Marco teórico

2.1. Arrays de micrófonos

2.2. Características acústicas de insectos xilófagos

2.3. Redes neuronales

Las redes neuronales artificiales (ANNs, *Artificial Neural Networks*) fueron inspiradas por sus análogas biológicas. Se tratan de modelos computacionales que permiten crear algoritmos complejos [5].

2.3.1. Redes neuronales tipo transformer

Un modelo transformer consiste en una red neuronal que es capaz de entender el contexto mediante un mecanismo denominado auto-atención. Esta arquitectura fue descrita por primera vez en un documento de Google en 2017, [Attention Is All You Need](#) [6]. Desde su aparición se han convertido en una herramienta clave en el ámbito del procesamiento del lenguaje natural (NLP, *Natural Language Processing*) así como ha supuesto la base de sistemas como Chat GPT (*Generative Pre-training Transformer*).

Esta arquitectura se presenta como una solución a los inconvenientes que tenían sus predecesores, las Redes neuronales recurrentes (RNN) y las Redes neuronales de memoria de corto-largo plazo (LSTM).

Figura 2.1.

Input Embedding

Las redes neuronales trabajan con valores numéricos, por tanto, es necesario convertir aquellas entradas que sean de un tipo de valor diferente en números. Esto se consigue mediante una capa de vectorización. Uno de los métodos más utilizados en el ámbito de las redes neuronales para representar palabras como vectores numéricos, se conoce como Word Embedding.

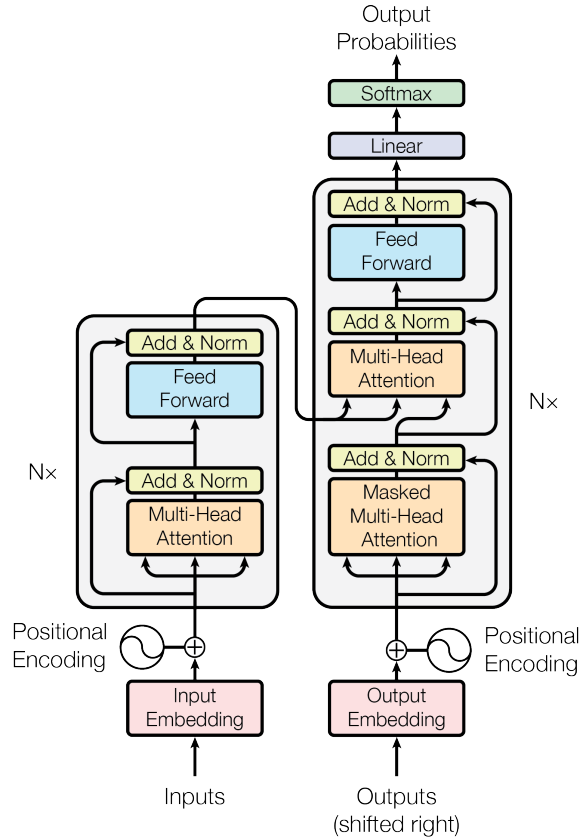


Figura 2.1: Arquitectura de un Transformer [6].

Positional Encoding

Self-Attention

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \quad (2.1)$$

Codificador

Decodificador

2.4. Herramientas de desarrollo

2.4.1. LabVIEW

2.4.2. Python

Capítulo 3

Procesado de señales

Para desarrollar un sistema capaz de **distinguir entre nudo y larva** es necesario capturar las señales así como su posterior procesamiento. Es esta segunda parte la que más decisiones involucra, puesto que al analizar las señales capturadas en bruto surgen numerosas formas de procesarlas cada una de ellas con sus ventajas y desventajas. Sin embargo, es una parte muy importante para el posterior desarrollo de un algoritmo de IA contar con una buena base de datos, con datos representativos de la realidad.

3.1. Adquisición de señales

Contar cómo se ha llevado a cabo la adquisición, condiciones ambientales, cuándo se capturaba, etc... (preguntar a AIF)

3.2. Segmentación de señales

Explicar cómo se ha pasado de la señal de 22ms a la de 4ms, qué parámetros se han tenido en cuenta para la selección, cómo ha sido el procesamiento, etc...

Explicar parámetros importantes sobre cómo aumentar o reducir el tiempo de las guardas afecta o no al resultado final tras la conformación

Como se muestra en la Figura 3.1 ...

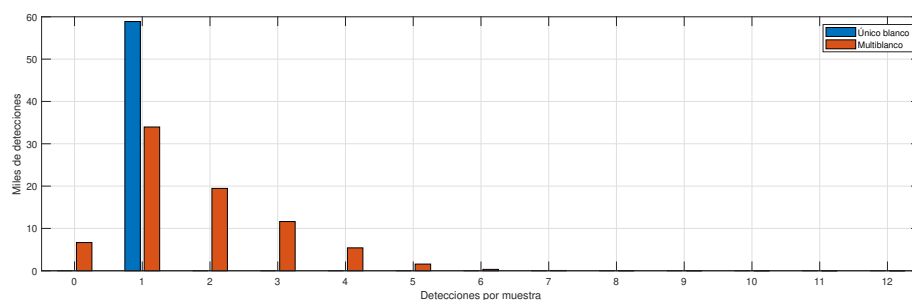


Figura 3.1: Número total de detecciones por captura con diferentes valores de guarda.

3.3. Conformación de las señales

Explicar qué métodos hay posibles (con FFT, con delay and sum, con retardos fraccionales...) y contar cómo se ha conformado las señales.

3.4. Creación de la base de datos (etiquetado)

Explicar qué parámetros se han usado para el etiquetado automático de las señales segmentadas.

3.5. Análisis de los datos

3.6. Cálculo de espectrogramas

Contar cómo se adecúan los datos crudos para el deep learning

Transformada wavelet continua en (3.1)

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t - \tau}{s} \right) dt \quad (3.1)$$

Transformada wavelet discreta en (3.2)

$$D[a, b] = \frac{1}{\sqrt{b}} \sum_{m=0}^{p-1} f[t_m] \psi^* \left[\frac{t_m - a}{b} \right] \quad (3.2)$$

Capítulo 4

Desarrollo de la red neuronal

Antes de diseñar una red neuronal compleja completa desde cero, es interesante buscar entre las redes existentes aquellas que tengan un propósito parecido y usen datos de entrenamiento similares a los de interés.

4.1. *Transfer learning*

4.2. *Fine tuning*

4.3. Diseño de las nuevas capas

4.4. Entrenamiento de la red

Capítulo 5

Resultados

Capítulo 6

Conclusiones

Bibliografía

- [1] R. D. Martínez, A. Izquierdo, J. J. Villacorta, L. del Val, and L. A. Basterra, “Acoustic detection and localisation system for hylotrupes bajulus l. larvae using a mems microphone array,” *Applied Acoustics*, vol. 213, p. 109618, 10 2023.
- [2] X. Liu, H. Zhang, Q. Jiang, L. Ren, Z. Chen, Y. Luo, and J. Li, “Acoustic denoising using artificial intelligence for wood-boring pests semanotus bifasciatus larvae early monitoring,” *Sensors 2022, Vol. 22, Page 3861*, vol. 22, p. 3861, 5 2022.
- [3] J. Schofield and D. Chesmore, “Automated acoustic identification of beetle larvae in imported goods using time domain analysis,” *Proceedings - European Conference on Noise Control*, pp. 5929–5934, 2008.
- [4] R. W. Mankin, A. Mizrach, A. Hetzroni, S. Levsky, Y. Nakache, and V. Soroker, “Temporal and spectral features of sounds of wood-boring beetle larvae: Identifiable patterns of activity enable improved discrimination from background noise,” *Florida Entomologist*, vol. 91, p. 241 – 248, 2008.
- [5] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O’Reilly Media, Inc, 2nd ed., 9 2019.
- [6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Łukasz Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in Neural Information Processing Systems*, vol. 2017-December, pp. 5999–6009, 6 2017.
- [7] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, “Transformers in time series: A survey,” *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2023-August, pp. 6778–6786, 2 2022.
- [8] T. Zhou, Z. Ma, Q. Wen, X. Wang, L. Sun, and R. Jin, “Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting,” *Proceedings of Machine Learning Research*, vol. 162, pp. 27268–27286, 1 2022.
- [9] F. A. Martín, “Desarrollo y análisis de clasificadores de señales de audio,” 10 2017.
- [10] P. Bilski, P. Bobiński, A. Krajewski, and P. Witomski, “Detection of wood boring insects’ larvae based on the acoustic signal analysis and the artificial intelligence algorithm,” *Archives of Acoustics*, vol. 42, pp. 61–70, 3 2017.
- [11] J. Schofield, “Real-time acoustic identification of invasive wood-boring beetles,” 2011.

- [12] X. Wei, M. Z. Hossain, and K. A. Ahmed, “A resnet attention model for classifying mosquitoes from wing-beating sounds,” *Scientific Reports 2022 12:1*, vol. 12, pp. 1–10, 6 2022.

Anexos