



UNIVERSIDAD DE VALLADOLID

ESCUELA TÉCNICA SUPERIOR DE  
INGENIEROS DE TELECOMUNICACIÓN

TRABAJO FIN DE Grado

GRADO EN INGENIERÍA DE TECNOLOGÍAS ESPECÍFICAS  
DE TELECOMUNICACIÓN

**APLICACIÓN DE TÉCNICAS DE DEEP LEARNING A  
SEÑALES DE PULSIOXIMETRÍA PARA LA DETECCIÓN  
AUTOMÁTICA DE LAS FASES DEL SUEÑO**

Autor:

**D. Daniel Sirgo Rodríguez**

Tutores:

**Dr. D. Fernando Vaquerizo Villar**

**Dr. D. Roberto Hornero Sánchez**

Valladolid, 7 de Julio de 2022



---

**TÍTULO:**                    **Aplicación de técnicas de *deep learning* a señales de pulsioximetría para la detección automática de las fases del sueño.**

**AUTOR:**                    **D. Daniel Sirgo Rodríguez**

**TUTORES:**                **Dr. D. Fernando Vaquerizo Villar  
Dr. D. Roberto Hornero Sánchez**

**DEPARTAMENTO:**        **Teoría de la Señal y Comunicaciones  
e Ingeniería Telemática**

---

**TRIBUNAL**

---

**PRESIDENTE:**            **Dr. D. Roberto Hornero Sánchez**

**SECRETARIO:**           **Dra. Dña. María García Gadañón**

**VOCAL:**                    **Dr. D. Jesús Poza Crespo**

**SUPLENTE 1:**            **Dr. D. Carlos Gómez Peña**

**SUPLENTE 2:**            **Dr. D. Mario Martínez Zarzuela**

---



# AGRADECIMIENTOS

Me gustaría agradecer, en primer lugar, a mis tutores Fernando Vaquerizo y Roberto Hornero por ofrecerme la oportunidad y experiencia de realizar un Trabajo Fin de Grado relacionado con un tema tan interesante como son las fases del sueño y permitirme conocer más sobre la investigación dentro de ese ámbito. En especial quiero agradecer a Fernando por resolverme todas las dudas y hacerme de guía en este trabajo. Gracias a él he aprendido mucho.

También mis agradecimientos a los miembros del Grupo de Ingeniería Biomédica con los que he compartido laboratorio.

Gracias a todos mis amigos y compañeros que me han acompañado durante mis años de carrera, haciendo que esta etapa haya sido muy bonita e inolvidable.

Por último, agradecer a mi madre por apoyarme cuando lo he necesitado y a mi novia, que me ha dado el ánimo y las ganas necesarias para poder terminar este trabajo de la mejor manera posible. Os quiero a las dos.

Muchas gracias a todos.



# RESUMEN DEL TFG

Las enfermedades del sueño en edad infantil son patologías muy relevantes en la sociedad actual debido a las consecuencias negativas que tienen para la salud y el desarrollo de los niños. Por tanto, es necesario un adecuado y objetivo diagnóstico y seguimiento de ellas. Para ello se vuelve imprescindible una buena detección de las fases del sueño.

En este trabajo fin de grado (TFG) se ha profundizado en la detección automática de las fases del sueño mediante técnicas de *deep learning*. Concretamente, se han usado redes de *deep learning* de dos tipos diferentes, redes neuronales convolucionales (CNNs) y CNNs combinadas con redes neuronales recurrentes (RNNs).

Como novedad en este TFG respecto a estudios previos se ha trabajado con señales de pulsioximetría, en lugar de señales más comunes como pueden ser el electroencefalograma (EEG) y el electrocardiograma (ECG). La adquisición del EEG y el ECG requiere de métodos costosos e intrusivos, y aunque consiguen buenos resultados, no resulta viable su implementación en entornos no supervisados, como puede ser el domicilio de los niños. En cambio, las señales de pulsioximetría (fotopletismografía: PPG, frecuencia de pulso: PR y saturación de oxígeno en sangre: SpO<sub>2</sub>) son fáciles de obtener mediante métodos mínimamente invasivos y de bajo coste en el domicilio de los niños. Por ello, el objetivo principal de este TFG es evaluar la utilidad de métodos de *deep learning* aplicados sobre señales de pulsioximetría para la detección automática de fases del sueño en niños con apnea obstructiva del sueño (AOS).

Para conseguir este objetivo se ha utilizado una base de datos compuesta por señales de pulsioximetría de 429 niños con AOS. Estas señales fueron preprocesadas por separado para realizar una correcta normalización y homogeneización. Posteriormente se ha realizado una revisión del estado del arte con estudios similares para una orientación de la metodología a aplicar. Se entrenaron dos arquitecturas de *deep learning* utilizando las señales de pulsioximetría para detectar las fases del sueño en cada época de 30 segundos de los registros del sueño: una red CNN y una red compuesta por una CNN combinada con una RNN (CNN+RNN). Los hiperparámetros de estas arquitecturas fueron optimizados para poder conseguir los mayores niveles de precisión posible.

Como resultados a destacar se ha obtenido una precisión de 5 clases del 68,09% y una Kappa de Cohen de 0,5778 con la arquitectura CNN+RNN entrenada con las señales de PR y SpO<sub>2</sub> combinadas, así como una precisión de 68,79% y un Kappa de 0,5870 con el modelo CNN+RNN obtenido a partir de la señal de PPG por separado. Estos resultados han sido superiores a los obtenidos con las arquitecturas CNN. Los resultados finales del trabajo muestran que son significativamente inferiores a los conseguidos en otros estudios con señales de EEG, pero mejores que en otros estudios que han utilizado señales de pulsioximetría. Por tanto, son unos resultados aceptables que nos ofrecen esperanza para poder continuar investigando y profundizando en el estudio de estas señales.

**Palabras clave:** Apnea obstructiva del sueño (AOS), *deep learning*, fases del sueño, fotopletismografía (PPG), frecuencia de pulso (PR), niños, pulsioximetría, redes neuronales convolucionales (CNN), redes neuronales recurrentes (RNN). saturación de oxígeno en sangre (SpO<sub>2</sub>).



# ABSTRACT

Sleep disorders in the pediatric population are pathologies very relevant in today's society due to their negative consequences for the health and development of the affected children. Therefore, a suitable and objective diagnosis and monitoring of them is necessary. In order to do so, it is essential to detect the sleep stages.

In this final degree project (*trabajo fin de grado*, TFG), we have studied in depth the automatic detection of sleep stages using deep learning techniques. Specifically, two different deep learning networks have been used: convolutional neural networks (CNNs) and CNNs combined with recurrent neural networks (RNNs).

As a novelty in this TFG with respect to previous studies, we have worked with pulse oximetry signals, instead of more common signals such as the electroencephalogram (EEG) and the electrocardiogram (ECG). EEG and ECG acquisition requires expensive and intrusive methods, and although they achieve good results, they are not feasible to implement in unsupervised environments, such as children's homes. In contrast, pulse oximetry signals (photoplethysmography: PPG, pulse rate: PR and blood oxygen saturation: SpO<sub>2</sub>) are easy to obtain by minimally invasive and low-cost methods in children's homes. Therefore, the main objective of this TFG is to evaluate the usefulness of deep learning methods applied on pulse oximetry signals for the automatic detection of sleep stages in children with obstructive sleep apnea (OSA).

To achieve this goal, a database of pulse oximetry signals from 429 children with OSA was used. These signals were preprocessed separately for correct normalization and homogenization. Subsequently, a state-of-the-art review was performed to provide guidance for the methodology to be applied. Two deep learning architectures were trained using pulse oximetry signals to detect sleep stages in each 30 seconds segment of the sleep recordings: a CNN network and a network composed of a CNN combined with a RNN (CNN+RNN). The hyperparameters of these architectures were optimized to achieve the highest possible levels of accuracy.

The CNN+RNN architecture fed with PR and SpO<sub>2</sub> signals achieved a 5-class accuracy of 68.09% and a Cohen's Kappa of 0.5778, whereas the CNN+RNN model obtained using the separate PPG signal reached an accuracy of 68.79% and a Kappa of 0.5870. These results were superior to those obtained with the CNN architecture. The final results of the work show that they are significantly lower than those obtained in other studies using EEG signals, but better than the reported by previous studies using pulse oximetry signals. Thus, these results are acceptable results and give us hope for further research and further study of automatic sleep staging from pulse oximetry signals.

**Keywords:** Children, Convolutional neural network (CNN), Deep Learning, Obstructive sleep apnea (OSA), Oxygen saturation (SpO<sub>2</sub>), Photoplethysmography (PPG), Pulse oximetry, Pulse Rate (PR), Recursive neural network (RNN), Sleep stages.



# ÍNDICE GENERAL

Capítulo 1: INTRODUCCIÓN .....	1
1.1. PROCESADO DE SEÑALES BIOMÉDICAS.....	1
1.1.1. ETAPAS DEL PROCESADO DE SEÑALES BIOMÉDICAS .....	1
1.1.2. TIPOS DE SEÑALES BIOMÉDICAS.....	1
1.2. APNEA OBSTRUCTIVA DEL SUEÑO INFANTIL .....	2
1.3. FASES DEL SUEÑO.....	2
1.4. PULSIOXIMETRÍA .....	4
1.5. DEEP LEARNING.....	6
1.6. HIPÓTESIS Y OBJETIVOS.....	7
1.7. ESTRUCTURA DEL TFG .....	8
Capítulo 2: DETECCIÓN AUTOMÁTICA DE LAS FASES DEL SUEÑO .....	9
2.1. APNEA DEL SUEÑO INFANTIL .....	9
2.2. CLASIFICACIÓN AUTOMÁTICA DE LAS FASES DEL SUEÑO.....	9
2.2.1. ENFOQUES CONVENCIONALES.....	10
2.2.2. PROCESADO AUTOMÁTICO DE SEÑALES BIOMÉDICAS.....	10
2.2.3. MÉTODOS DE DEEP-LEARNING .....	12
2.3. COMPARACIÓN Y ELECCIÓN DEL MÉTODO A IMPLEMENTAR .....	12
Capítulo 3: SUJETOS Y SEÑALES .....	15
3.1. BASE DE DATOS .....	15
3.1.1. POBLACIÓN BAJO ESTUDIO.....	15
3.1.2. CARACTERÍSTICAS DE LAS SEÑALES DE PULSIOXIMETRÍA.....	16
3.2. DIVISIÓN EN ENTRENAMIENTO, VALIDACIÓN Y TEST .....	16
Capítulo 4: METODOLOGÍA .....	19
4.1. REDES NEURONALES CONVOLUCIONALES.....	19
4.1.1. ARQUITECTURA CNN APLICADA.....	21
4.2. REDES NEURONALES RECURRENTES .....	24
4.2.1. ARQUITECTURA CNN+RNN APLICADA .....	24
4.3. ANÁLISIS ESTADÍSTICO .....	25
4.3.1. MÉTRICAS DE RENDIMIENTO DE LA CLASIFICACIÓN DE FASES DEL SUEÑO .....	25
4.3.2. MÉTRICAS DE RENDIMIENTO DE LA ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS.....	26

Capítulo 5: RESULTADOS.....	27
5.1. OPTIMIZACIÓN DE HIPERPARÁMETROS.....	27
5.1.1. OPTIMIZACIÓN DE LA ARQUITECTURA CNN .....	27
5.1.2. OPTIMIZACIÓN DE LA ARQUITECTURA CNN+RNN .....	28
5.2. CLASIFICACIÓN DE FASES DEL SUEÑO.....	28
5.2.1 CLASIFICACIÓN DE FASES DEL SUEÑO DE LA ARQUITECTURA CNN.....	29
5.2.2 CLASIFICACIÓN DE FASES DEL SUEÑO DE LA ARQUITECTURA CNN+RNN.....	30
5.3. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS .....	31
5.3.1. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS A PARTIR DE LA ARQUITECTURA CNN.....	31
5.3.2. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS A PARTIR DE LA ARQUITECTURA CNN+RNN.....	34
Capítulo 6: DISCUSIÓN .....	37
6.1. CLASIFICACIÓN DE FASES DEL SUEÑO.....	37
6.2. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS .....	38
6.3. COMPARACIÓN CON OTROS ESTUDIOS.....	39
6.4. LIMITACIONES .....	41
Capítulo 7: CONCLUSIONES Y LÍNEAS FUTURAS .....	43
7.1. CONTRIBUCIONES .....	43
7.2. CONCLUSIONES .....	43
7.3. LÍNEAS FUTURAS .....	44

# ÍNDICE DE FIGURAS

---

## CAPÍTULO 1: INTRODUCCIÓN

---

Figura 1.1: Intervalo P-P de una señal PPG.....	5
Figura 1.2: Ejemplo de una señal de PPG de la base de datos Childhood Adenotonsillectomy Trial (CHAT).....	5
Figura 1.3: Ejemplo de una señal de PR de la base de datos CHAT.....	6
Figura 1.4: Ejemplo de una señal de SpO <sub>2</sub> de la base de datos CHAT.....	6
Figura 1.5: Diferencias entre <i>Machine Learning</i> y <i>Deep Learning</i> .....	7

## CAPÍTULO 4: METODOLOGÍA

---

Figura 4.1: Diagrama de Inteligencia Artificial, <i>Machine Learning</i> y <i>Deep Learning</i> .....	19
Figura 4.2: Ejemplo de conexión entre neuronas en una CNN.....	19
Figura 4.3: Diferencias entre <i>Average Pooling</i> y <i>Max Pooling</i> .....	20
Figura 4.4: Ejemplo de convolución 2-D.....	23
Figura 4.5: Ejemplo de una red neuronal sin dropout y con dropout.....	23
Figura 4.6: Ejemplo de recursividad en una RNN.....	25

## CAPÍTULO 5: RESULTADOS

---

Figura 5.1: Matrices de confusión de las señales PR, SpO <sub>2</sub> Y PR+ SpO <sub>2</sub> obtenidas con CNN.....	29
Figura 5.2: Matrices de confusión de las señales PPG y PPG+ SpO <sub>2</sub> obtenidas con CNN.....	29
Figura 5.3: Matrices de confusión de las señales PR, SpO <sub>2</sub> Y PR+ SpO <sub>2</sub> obtenidas con CNN+RNN.....	30
Figura 5.4: Matrices de confusión de las señales PPG y PPG+ SpO <sub>2</sub> obtenidas con CNN+RNN.....	30

Figura 5.5: Gráfico de dispersión de las señales PR, SpO <sub>2</sub> y PR+ SpO <sub>2</sub> para el TST obtenidos con CNN. ....	33
Figura 5.6: Gráfico de dispersión de las señales PPG y PPG+ SpO <sub>2</sub> para el TST obtenidos con CNN. ....	33
Figura 5.7: Gráfico de dispersión de las señales PR, SpO <sub>2</sub> y PR+SpO <sub>2</sub> para el TST obtenidos con CNN + RNN. ....	35
Figura 5.8: Gráfico de dispersión de las señales PPG y PPG+SpO <sub>2</sub> para el TST obtenidos con CNN + RNN. ....	36

# ÍNDICE DE TABLAS

---

## CAPÍTULO 1: INTRODUCCIÓN

---

Tabla 1.1: Fases del sueño. ....	4
----------------------------------	---

## CAPÍTULO 3: SUJETOS Y SEÑALES

---

Tabla 3.1: Características de la base de datos para las señales de PR+ SpO <sub>2</sub> . ....	15
--	----

Tabla 3.2: Características de la base de datos para las señales de PPG+ SpO <sub>2</sub> . ....	15
---	----

Tabla 3.3: División de señales para entrenamiento, validación y test de las combinaciones PR, SpO <sub>2</sub> y PR+SpO <sub>2</sub> .....	17
--	----

Tabla 3.4: División de señales para entrenamiento, validación y test de las combinaciones PPG y PPG+SpO <sub>2</sub> .....	17
--	----

## CAPÍTULO 5: RESULTADOS

---

Tabla 5.1: Resultados de precisión obtenidos con las diferentes etapas de solapamiento....	27
--	----

Tabla 5.2: Resultados de precisión obtenidos con los diferentes valores de dropout y bloques convolucionales. ....	28
--	----

Tabla 5.3: Resultados de precisión obtenidos con los diferentes tamaños de secuencias y unidades GRU. ....	28
--	----

Tabla 5.4: Resultados de precisión, Kappa, MF1 y <i>per-class F1-score</i> en test para la estructura CNN. ....	29
---	----

Tabla 5.5: Resultados de precisión, Kappa, MF1 y <i>per-class F1-score</i> en test para la estructura CNN+RNN. ....	30
---	----

Tabla 5.6: Resultados del p-valor, ICC, error medio e índice de correlación para la fase N1 para la estructura CNN.....	31
---	----

Tabla 5.7: Resultados del p-valor, ICC, error medio e índice de correlación para la fase N2 para la estructura CNN.....	31
---	----

Tabla 5.8: Resultados del p-valor, ICC, error medio e índice de correlación para la fase N3 para la estructura CNN.....	31
---	----

Tabla 5.9: Resultados del p-valor, ICC, error medio e índice de correlación para la fase REM para la estructura CNN.....	32
Tabla 5.10: Resultados del ICC y error medio para la fase W para la estructura CNN.....	32
Tabla 5.11: Resultados del p-valor, ICC, error medio e índice de correlación para la fase N1 para la estructura CNN+RNN.....	34
Tabla 5.12: Resultados del p-valor, ICC, error medio e índice de correlación para la fase N2 para la estructura CNN+RNN.....	34
Tabla 5.13: Resultados del p-valor, ICC, error medio e índice de correlación para la fase N3 para la estructura CNN+RNN.....	34
Tabla 5.14: Resultados del p-valor, ICC, error medio e índice de correlación para la fase REM para la estructura CNN+RNN.....	34
Tabla 5.15: Resultados del ICC y error medio para la fase W para la estructura CNN+RNN.....	35

## CAPÍTULO 6: DISCUSIÓN

---

Tabla 6.1: Estudios consultados en la revisión del estado del arte.....	40
---	----

# Capítulo 1: INTRODUCCIÓN

---

## 1.1. PROCESADO DE SEÑALES BIOMÉDICAS

Una señal biomédica es una señal procedente del cuerpo humano que nos permite saber diversas características y el estado en el que se encuentra una persona. Tiene gran importancia porque nos proporciona información de ayuda al diagnóstico y monitorización de numerosas patologías. Su procesado constituye un ámbito muy extenso de investigación en el que diversas disciplinas son englobadas para analizar las señales generadas por el cuerpo humano, en busca de posibles alteraciones o anomalías en su correcto funcionamiento [1].

El principal objetivo de esta área es ayudar a los médicos a diagnosticar enfermedades facilitándoles el análisis y ofreciéndoles señales con la más alta precisión y calidad posibles [2]. Se busca, además de un correcto diagnóstico, un seguimiento fiable de la evolución del paciente. En los últimos años se han conseguido grandes avances tecnológicos, lo que ha permitido una mejora significativa en las técnicas de la ingeniería biomédica, con métodos más baratos, eficientes y prácticos.

La Ingeniería Biomédica es una disciplina que utiliza diversas técnicas de ingeniería y las aplica a problemas relacionados con la medicina y la biología para realizar un análisis y resolución de ellos. En este terreno, donde se une la medicina con la ingeniería, se busca ganar objetividad para facilitar una opinión de un profesional de la medicina, en ningún caso se busca su sustitución. Muchas veces los diagnósticos médicos se basan en hipótesis, creencias o en la propia experiencia del profesional sanitario que nos está realizando el análisis y gracias a las técnicas o algoritmos de la Ingeniería Biomédica conseguimos que esos diagnósticos estén dotados de la mayor objetividad posible [1].

Estas técnicas no solo sirven para mejorar los diagnósticos directamente, sino que también nos permiten medir la magnitud de las señales de la mejor forma posible, reduciendo al mínimo posible interferencias y ruido que nos puedan alterar la información contenida en ellas.

### 1.1.1. ETAPAS DEL PROCESADO DE SEÑALES BIOMÉDICAS

Toda señal biomédica debe ser procesada correctamente. Para ello se necesita realizar las siguientes etapas [3]:

- Obtención y registro: Adquisición, muestreo, cuantificación, digitalización y cualquier preprocesado necesario. En este paso se pueden utilizar diferentes técnicas, invasivas o no invasivas. Las señales utilizadas en este TFG se han obtenido mediante una técnica no invasiva, la pulsioximetría, que se basa en las propiedades ópticas de la hemoglobina.
- Procesado: Segmentación, filtrado y detección de patrones deseados.
- Clasificación: Obtención de características y clasificación de la señal.

### 1.1.2. TIPOS DE SEÑALES BIOMÉDICAS

Dentro de las señales biomédicas podemos encontrarnos con diferentes categorías en función de diferentes criterios como pueden ser su adquisición, su origen o sus características principales. También pueden ser deterministas o aleatorias en función de si pueden ser

predichas o descritas matemáticamente. Las deterministas nos permiten calcular valores fiables en función de las ya obtenidos. En cambio, las no deterministas precisan de un análisis estadístico. En la práctica, las señales biomédicas mezclan los dos tipos de señales y suelen ser continuas en el tiempo y de pequeña amplitud [1].

Una posible clasificación en función de sus características podría ser:

- Señales acústicas [1]: Producidas por procesos biológicos que generan sonido como la circulación sanguínea o la respiración. Son fáciles de obtener con métodos poco invasivos como los micrófonos.
- Señales eléctricas: Producidas por potenciales de las membranas de las células nerviosas y musculares como el electroencefalograma. Se obtienen con métodos un poco invasivos con uso de electrodos colocados en distintas zonas del cuerpo del paciente.
- Señales de bioimpedancias [1]: Producidas por la impedancia de los tejidos que aportan mucha información de su composición y de su actividad.
- Señales mecánicas [1]: Producidas por procesos mecánicos que generan desplazamientos, fuerza y tensión. Su obtención precisa de métodos invasivos.
- Señales ópticas [1]: Producidas por el sistema biológico. En este TFG se han utilizado las señales de fotopleletismografía (PPG), saturación de oxígeno en sangre ( $SpO_2$ ) y frecuencia de pulso (PR), adquiridas mediante un pulsioxímetro, colocado en el dedo o lóbulo de la oreja del paciente.
- Señales magnéticas [1]: Producidas por los campos magnéticos de los órganos. Son muy débiles y su obtención es compleja y costosa.
- Señales químicas [1]: Producidas al realizar pruebas en laboratorios con tejidos orgánicos.

## 1.2. APNEA OBSTRUCTIVA DEL SUEÑO INFANTIL

La apnea obstructiva del sueño (AOS) es un trastorno respiratorio caracterizado por la recurrencia de eventos de apnea e hipopnea durante el sueño, que derivan en una arquitectura del sueño fragmentada y no reparadora [4]. En caso de no ser tratada, la AOS puede originar consecuencias muy negativas para la salud de los niños, como complicaciones en el crecimiento, en el desarrollo cognitivo y en el comportamiento [5]. Por tanto, la detección temprana de la AOS infantil es un asunto de suma importancia, lo cual incluye la detección de las fases del sueño en niños que padecen AOS. La prueba de diagnóstico estándar de la AOS es la polisomnografía nocturna, que es un método costoso, complejo, altamente intrusivo para los niños y presenta disponibilidad limitada, lo que limita su efectividad [6], [7]. Las principales alternativas se centran en analizar un conjunto reducido de señales, entre las que destacan las señales de pulsioximetría por su sencillez y posibilidad de ser registradas en el propio domicilio del paciente mediante equipos portátiles de menor coste [8].

## 1.3. FASES DEL SUEÑO

El sueño es un estado físico y mental en el que el cerebro, músculos, ojos, corazón y pulmones producen gran actividad [9]. Durante el sueño se llevan a cabo procesos fisiológicos como el metabolismo y la consolidación de la memoria [10]. Es, sin duda, uno de los principales factores para tener un correcto funcionamiento a lo largo del día. Por desgracia, el estrés producido en la sociedad actual, las nuevas tecnologías, entre otros factores, han producido que la calidad del sueño de la población empeore, generando diversos problemas. Asimismo, diversos trastornos

y enfermedades (como la AOS descrita anteriormente) afectan a la calidad del sueño de los niños, destacando así su importancia a la hora de diagnosticarlos [10].

Como se he comentado anteriormente, un diagnóstico temprano de la AOS y un posterior tratamiento requieren detección de las fases del sueño. La prueba de diagnóstico estándar de la AOS es la polisomnografía nocturna, que consiste en la monitorización del sueño mediante el registro de múltiples señales neurofisiológicas y cardiorrespiratorias. A partir de estas señales, especialistas médicos siguen las reglas de la Academia Americana de Medicina del Sueño (*American Academy of Sleep Medicine, AASM*) [11] para anotar las fases del sueño y los eventos respiratorios (apneas e hipopneas).

De acuerdo con la AASM, el sueño se divide en varias fases, dos de ellas muy diferenciadas, la fase REM (*Rapid Eye Movement*) en la que se producen movimientos oculares rápidos y la fase no-REM (NREM), conocida como sueño de ondas lentas. En una clasificación más profunda, la NREM está compuesta por cuatro estados que se relacionan con el grado de profundidad del sueño. Nos encontramos con los estados N1 Y N2, conocidos como sueño ligero, y los estados N3 Y N4, conocidos como sueño profundo [12]. En nuestro trabajo, al igual que en la gran mayoría de estudios, las fases de sueño profundo las unimos en una única fase. Necesitamos conocer los patrones de cada fase y estados para detectar las posibles anomalías y patologías asociadas.

A lo largo del sueño se pasan varias veces por las diversas fases, lo cual se conoce como ciclos de sueño. Durante los ciclos del sueño nos encontramos con diferentes transiciones entre las fases del sueño. Por ejemplo, es común encontrarse con transiciones de las fases N1-N2-N3 o al revés [12]. Sin embargo, pasar de estar despierto a sueño profundo es complicado y menos probable.

La detección de estas fases del sueño la realizan los técnicos mediante el análisis visual de las señales de electroencefalografía (EEG), electrooculografía (EOG) y electromiograma submentoniano (EMG), para así clasificar cada época de 30-s de sueño en una de las siguientes 5 fases: despierto (Wake, W), niveles 1 (N1), 2 (N2) y 3 (N3) de sueño NREM y sueño REM [11], [12], como se muestra en la tabla 1.1. Estas anotaciones son utilizadas, junto con las de los eventos de apnea e hipopnea para obtener el índice de apnea-hipopnea (AHI: número de apneas e hipopneas por hora de sueño), que es la variable utilizada en la clínica para el diagnóstico de la AOS infantil.

A pesar de su efectividad, la PSG es un método costoso, complejo, altamente intrusivo para los niños y presenta disponibilidad limitada, lo que retrasa el diagnóstico y tratamiento de los niños afectados [6][7]. En este sentido, la anotación de las fases del sueño lleva una cantidad considerable de tiempo (en torno a las 2 horas) y presenta una gran variabilidad entre distintos especialistas, lo cual puede afectar al diagnóstico de los niños afectados [13]. Por ello, se han propuesto distintas alternativas centradas en el análisis automático de un conjunto reducido de señales. En este sentido, un enfoque habitual es el análisis automático de un canal (o varios) de EEG, lo que conlleva un método complejo que en cierto grupo de personas es especialmente incómodo, como en el caso de trabajo realizado, que tratamos con señales de niños.

En este TFG se han utilizado las señales de pulsioximetría, que destacan por ser adquiridas de manera no invasiva, así como por su sencillez y posibilidad de ser registradas en el propio domicilio del paciente mediante equipos portátiles de bajo coste [8].

**Tabla 1.1: Fases del sueño.**

FASE	PROFUNDIDAD DEL SUEÑO
W (Wake)	Despierto
N1	Somnolencia
N2	Sueño ligero
N3	Sueño profundo
REM	Movimiento rápido de ojos

En nuestro trabajo tratamos con señales de una base de datos de niños, que como en la mayoría de casos, difieren de las señales de los adultos. Esta diferencia es notable si trabajamos con EEG o con señales cardíacas y respiratorias.

La detección de estas fases se busca realizarla de forma automática. Para ello, diversas propuestas han sido publicadas y son conocidas como ASSC (*Automatic Sleep Stage Classification*) [11]. En nuestro trabajo vamos a emplear técnicas de *deep learning* para la detección automática de las fases del sueño.

## 1.4. PULSIOXIMETRÍA

La pulsioximetría es una técnica no invasiva que permite determinar el porcentaje de saturación de oxígeno de la hemoglobina en la sangre ( $SpO_2$ ) con ayuda de métodos fotoeléctricos, además de la frecuencia cardíaca. Su uso es muy útil en diagnóstico de enfermedades respiratorias y cardíacas, antes de someterse a una operación y para detectar trastornos del sueño, como la apnea. En este trabajo, en concreto, hemos utilizado tres tipos de señales obtenidas mediante esta técnica: PPG,  $SpO_2$  Y PR [14].

La señal PPG es una señal óptica no invasiva empleada en la detección de cambios del flujo sanguíneo en el tejido micro vascular. El principio de obtención se basa en que la sangre arterial presenta una absorción de la luz diferente al resto de tejidos en ciertas longitudes de onda, debido a la fluctuación de las proteínas y la hemoglobina causada por la circulación de sangre oxigenada. Cuando el corazón bombea la sangre el flujo sanguíneo que llega a los capilares cercanos a la superficie de la piel aumenta, esto supone una mayor absorción de la luz. En cambio, cuando la sangre vuelve a corazón el flujo disminuye y la absorción de luz disminuye también [14].

La fotopletismografía (PPG) utiliza una fuente de luz y un fotodetector en la superficie de la piel para medir las variaciones volumétricas de la circulación sanguínea [15].

La señal PPG comprende componentes pulsátiles (AC) y superpuestos (DC). El componente AC lo proporcionan las variaciones sincrónicas del volumen sanguíneo que surgen de los latidos del corazón. El componente DC está formado por la respiración, la actividad del sistema nervioso simpático y la termorregulación. El componente AC representa los cambios en el volumen sanguíneo, que son causados por la actividad cardíaca y dependen de las fases sistólica y diastólica. La fase sistólica (también llamada "tiempo de subida") comienza con un valle y termina con el pico sistólico de la onda de pulso. El final de la onda de pulso está marcado por otro valle al final de la fase diastólica [15].

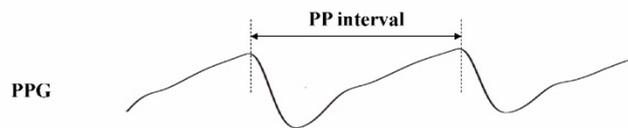
La señal PPG puede utilizarse para medir las variaciones entre los intervalos de tiempo de los latidos del corazón (intervalo pico a pico o P-P), como observamos en la figura 1.1. La variación puede deberse a muchos factores, como la edad del individuo, las condiciones cardíacas y el

estado físico. De ese intervalo P-P obtenemos la señal *Pulse Rate* (PR) utilizada en este trabajo. La señal PR es una señal derivada de la variación de la frecuencia cardiaca. En la imagen podemos observar la forma de una señal PR obtenida a partir de la señal de PPG [15].

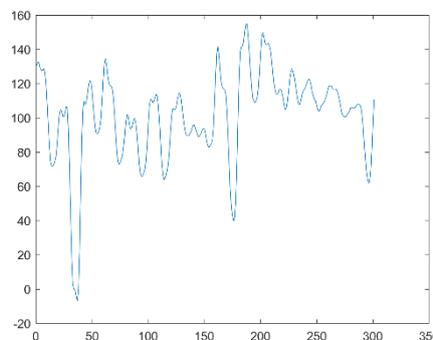
Por último, a partir de la señal PPG podemos obtener el nivel de saturación en sangre. La señal SpO<sub>2</sub> (saturación de oxígeno) permite el cálculo del porcentaje de oxígeno en sangre. En condiciones normales la señal SpO<sub>2</sub> esta correlacionada con el ritmo cardiaco del paciente y su valor máximo está en fase con la onda R de la señal ECG [16]. El nivel de oxígeno en sangre normal es de aproximadamente 75 a 100 mm HG. En un pulsioxímetro los valores suelen oscilar entre el 95 y el 100%, siendo los valores inferiores a 90% bajos muy posiblemente debidos a alguna patología [16].

La manera de obtención de estas es más compleja que la de PR. El nivel de saturación de oxígeno (SpO<sub>2</sub>) se calcula tomando la relación de los componentes de AC y DC de la señal del fotopletismógrafo (PPG) medida por el pulsioxímetro a dos longitudes de onda diferentes, una, alrededor de 660 nm (espectro rojo) y otra, alrededor de 940 nm (espectro infrarrojo). Para una estimación precisa de la SpO<sub>2</sub>, es necesario extraer el componente de AC de la señal mediante el procesamiento de la misma, donde la detección precisa de los picos es un elemento crucial y difícil. Los picos de la señal PPG deben ser identificados correctamente para estimar con precisión la SpO<sub>2</sub>. Los falsos picos y valles introducen grandes errores en el cálculo debido a la estimación incorrecta de las amplitudes, especialmente porque los picos de la forma de onda PPG pueden consistir menos del 1% de la señal. Por lo tanto, se requiere un filtrado para extraer la amplitud de AC correcta [16].

En el trabajo realizado analizamos las señales de PPG (se representa en la figura 1.2), PR (se representa en la figura 1.3) y SpO<sub>2</sub> (se representa en la figura 1.4). Las señales son mostradas en un mismo periodo de tiempo para comprobar sus diferencias.



**Figura 1.1:Intervalo P-P de una señal PPG**



**Figura 1.2:Ejemplo de una señal de PPG de la base de datos Childhood Adenotonsillectomy Trial (CHAT).**

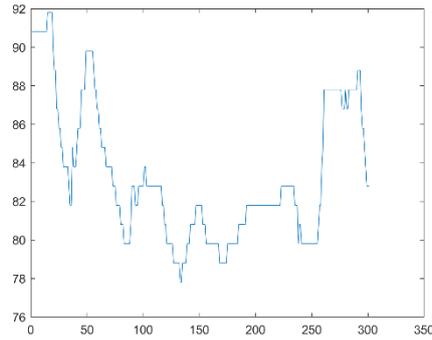


Figura 1.3: Ejemplo de una señal de PR de la base de datos CHAT.

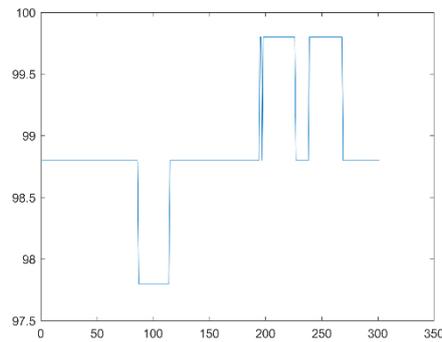


Figura 1.4: Ejemplo de una señal de SpO2 de la base de datos CHAT.

## 1.5. DEEP LEARNING

El *deep learning* es un conjunto de algoritmos de aprendizaje automático (*machine learning*) que modela abstracciones de alto nivel en datos usando arquitecturas computacionales con transformaciones no lineales e iterativas [17].

Este concepto viene derivado de la inteligencia artificial, una disciplina que intenta replicar y desarrollar inteligencia a través de máquinas y ordenadores. De esta ciencia proviene el *machine learning*, una rama cuyo objetivo es desarrollar técnicas para que las computadoras aprendan. El *deep learning* es *machine learning* pero no todo el *machine learning* es *deep learning*. La principal diferencia es que el *deep learning* es capaz de procesar datos crudos sin procesar (*raw data*) mientras que el *machine learning* en general conlleva una extracción previa de las características donde se puede perder cierta información. Estas diferencias se muestran de forma gráfica en la figura 1.5 [18].

El *deep learning* lleva a cabo procesos de *machine learning* usando redes neuronales artificiales compuestas de varios niveles jerárquicos, que permiten extraer patrones más complejos que los métodos convencionales. El *deep learning* es un sistema con mucho potencial ya que puede aplicarse a grandes volúmenes de datos con tasa de éxito elevada [17].

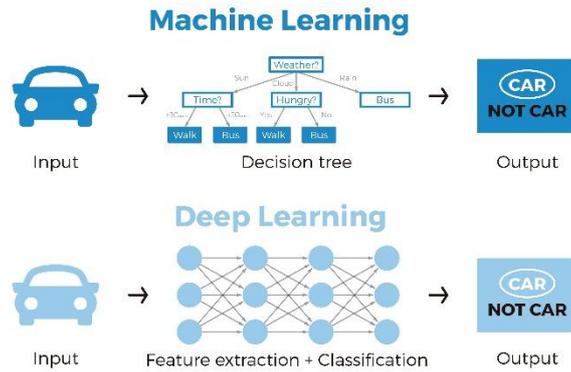


Figura 1.5: Diferencias entre Machine Learning y Deep Learning.

Para ejecutar los algoritmos, normalmente se requiere de una computación de propósito general en unidades de procesamiento gráfico (GPU). Mediante este procedimiento podemos entrenar a nuestras redes neuronales y posteriormente poder evaluar nuestros modelos creados [17].

Algunos de los usos más extendidos pueden ser el reconocimiento de voz, identificación de imágenes y caras, monitorización en tiempo real, realización de predicciones o el procesado de lenguaje y de señales [17]. En nuestro trabajo lo vamos a emplear para la detección automática de las fases del sueño, partiendo de unas señales con las fases ya identificadas para su entrenamiento.

## 1.6. HIPÓTESIS Y OBJETIVOS

En el TFG realizado se parte de la hipótesis de que los métodos de *deep learning* pueden ser de ayuda a la detección automática de las fases del sueño a partir de señales de pulsioximetría. En este sentido, se pretenden conseguir resultados aceptables pero conscientes de que posiblemente sean peores que con otras señales de partida como el EEG.

El objetivo final a largo plazo es obtener un método automático, no invasivo, económico y fácil de usar para analizar la arquitectura del sueño en la ayuda al diagnóstico de la AOS infantil, así como otras anomalías y patologías relacionadas con el sueño. Este TFG es un punto de partida para conseguir ese objetivo. Para conseguir detectar las fases del sueño, se emplea algoritmos de detección automática basados en técnicas de *deep learning*. Estos algoritmos son aplicados a señales de pulsioximetría de una base de datos de niños que proviene del estudio *The Childhood Adenotonsillectomy Trial* (CHAT) [19].

Los objetivos específicos del trabajo son indicados a continuación:

1. Revisar el estado del arte sobre diferentes algoritmos, técnicas y señales publicados en revistas científicas relacionados con la detección automática de las fases del sueño.
2. Preprocesar las señales de pulsioximetría de la base de datos de CHAT.
3. Diseñar y aplicar las técnicas de *deep learning* seleccionadas, *Convolutional Neural Networks* (CNNs) y *Recursive Neural Networks* (RNNs), a las señales ya preprocesadas.
4. Analizar los resultados obtenidos.
5. Implementar mejoras a las redes neuronales para mejorar los resultados.

6. Extraer las conclusiones finales del estudio.

Para la realización se ha seguido el siguiente plan de trabajo:

- Fase 1: Realización de un curso de Python para introducción en *Deep Learning* y redes neuronales [20].
- Fase 2: Revisión del estado del arte.
- Fase 3: Preprocesado de las señales de pulsioximetría de la base de datos.
- Fase 4: Diseño de las metodologías de *deep learning* y análisis de resultados.
- Fase 5: Modificación de la metodología en busca de mejoras en los resultados.
- Fase 6: Análisis de los resultados finales.
- Fase 7: Extracción de las conclusiones y realización de la memoria del Trabajo Fin del Grado (TFG).

## 1.7. ESTRUCTURA DEL TFG

El Trabajo Fin de Grado realizado se divide en 7 capítulos:

- INTRODUCCIÓN: En este capítulo se explica brevemente en qué consiste la ingeniería biomédica, la obtención de señales, las señales que se emplean en el trabajo, la hipótesis y objetivos del estudio, así como la estructura del trabajo.
- DETECCIÓN AUTOMÁTICA DE LAS FASES DEL SUEÑO: En este capítulo se explican conceptos claves de las fases del sueño y de su detección.
- BASE DE DATOS: En este capítulo se muestra la base de datos, los sujetos y señales empleados.
- METODOLOGÍA: En este capítulo se explica la metodología empleada y el funcionamiento y estructura de las arquitecturas de *deep learning*. Primeramente, las características generales de ese tipo de redes y, por último, se centra en el modelo diseñado y empleado.
- RESULTADOS: En este capítulo se exponen los resultados obtenidos y la clasificación conseguida.
- DISCUSIÓN: En este capítulo se realiza la discusión sobre los resultados obtenidos y se analiza la viabilidad del método empleado.
- CONCLUSIONES Y LÍNEAS FUTURAS: En este capítulo se finaliza el trabajo mediante las conclusiones obtenidas y las posibilidades futuras del método y su viabilidad.

Por último, tras los siete capítulos se incluye la bibliografía empleada en el Trabajo Fin de Grado.

# Capítulo 2: DETECCIÓN AUTOMÁTICA DE LAS FASES DEL SUEÑO

---

## 2.1. APNEA DEL SUEÑO INFANTIL

La apnea obstructiva del sueño (AOS) es una entidad clínica bien conocida en los adultos pero que ha sido menos estudiada en la población infantil. En diferentes estudios se ha demostrado que el AOS infantil tiene una alta prevalencia (1-5%) y que sus consecuencias pueden ser importantes para la salud de los niños, así como afectar gravemente a su desarrollo [21]. Por tanto, el diagnóstico temprano de la AOS infantil se ha convertido en un asunto de gran importancia, el cual incluye la caracterización de las fases del sueño.

Como se ha comentado anteriormente, el período de sueño se divide en dos grandes fases: la REM y la NREM. La fase NREM se subdivide a su vez en ligero (N1 y N2) y profundo (N3-N4).

El ciclo de sueño empieza por la fase NREM (pasando de ligero a profundo) y acaba por la fase REM. Estos ciclos se suceden de forma consecutiva durante tres o cuatro veces por la noche en sujetos normales. Si el sueño es interrumpido y hay un despertar, los ciclos del sueño no se cumplen, por lo que el sueño no es reparador [22].

En el período NREM el tono muscular de la vía aérea superior disminuye, pero en el sueño REM esta disminución del tono muscular se hace máxima, lo cual facilita que las paredes de la faringe tiendan a colapsarse y se favorece el cierre total o parcial de la vía aérea, lo que se conoce como apnea. Una apnea es el cese del flujo de aire por la nariz o la boca. Una hipopnea es una disminución del flujo de aire que entra por la nariz o la boca, que causa una caída en la saturación del oxígeno y/o termina en un despertar transitorio no consciente [22].

Así, las apneas y/o hipopneas acontecidas en el AOS son consecuencia de la obstrucción de las vías aéreas altas durante la noche. Esta obstrucción se localiza concretamente en la faringe y es el resultado de un desequilibrio entre las fuerzas que sirven para dilatar la faringe y las que favorecen su obstrucción [22].

Durante la noche el tono muscular de la vía aérea se halla reducido y, por tanto, se favorece el colapso. El diafragma que se mantiene activo va a tener que luchar contra la resistencia aumentada de la vía aérea superior y en este esfuerzo es cuando se puede producir el microdespertar, que va a coincidir con la apertura de la vía respiratoria y la normalización de la respiración. La sucesión de microdespertares durante la noche provoca la desestructuración del ciclo del sueño y que éste no sea reparador [22].

Por todo lo anterior, es necesario conocer la fase del sueño donde se producen los sucesos.

## 2.2. CLASIFICACIÓN AUTOMÁTICA DE LAS FASES DEL SUEÑO

Los métodos de detección automática de las fases del sueño, gracias al avance en las técnicas

de procesamiento de señales biomédicas, han obtenido las capacidades suficientes para superar los resultados de la clasificación manual [23].

### 2.2.1. ENFOQUES CONVENCIONALES

Los enfoques tradicionales se basan en técnicas basadas en el análisis visual para detectar las fases en la señal de EEG, diferenciando las distintas bandas de frecuencias y las formas de onda más comunes durante del sueño, mediante software [24]. Inicialmente, se solía diferenciar únicamente entre sueño ligero y sueño profundo, utilizando para ello algoritmos de detección de los pasos de la señal por cero y detección de valores máximos para la representación gráfica [14].

En 1989 se desarrolló la máquina de estados, el primer método de detección automática de las fases del sueño. En ella, cada estado representa una fase del sueño y las transiciones se realizaban aplicando la teoría de la evidencia de Dempster-Shafer [25].

La clasificación visual de las fases del sueño se puede realizar siguiendo las reglas establecidas por distintas organizaciones, siendo la más utilizada actualmente la AASM [11].

El desarrollo de algoritmos de detección automática de fases del sueño requiere de un elevado número de señales biomédicas provenientes de diferentes sujetos. En muchos estudios, estas señales se consiguen utilizando bases de datos ya existentes. Cada base de datos indica el número de sujetos, sus características, la frecuencia de muestreo aplicada y la técnica de adquisición de las señales.

Las señales biomédicas pueden contener diferentes artefactos que alteran la información que contienen y su posterior análisis. Por lo que es necesaria una etapa inicial de preprocesado que los atenúe o elimine del ruido.

### 2.2.2. PROCESADO AUTOMÁTICO DE SEÑALES BIOMÉDICAS

El procesamiento automático de señales biomédicas permite la detección automática de las fases del sueño, con las ventajas que ello conlleva respecto al análisis visual.

Como se he comentado, las señales biomédicas pueden contener diferentes artefactos que lleven a una mala interpretación de las mismas. Por ello, es necesaria una etapa inicial de preprocesado que los atenúe o elimine para dar más importancia a la componente informativa de la señal. Suelen ser métodos de filtrado digital, regresión (obteniendo un modelo que imite el comportamiento del ruido), transformadas (como la *Wavelet Transform* (WT)) o separación ciega de fuentes (separando el ruido basándose en que proviene de diferente fuente a la señal original) [10].

Una vez preprocesadas, el procesamiento de estas señales se suele realizar en 3 fases [10]:

- Extracción de características, donde se obtienen parámetros de la señal biomédica mediante técnicas de análisis en el dominio del tiempo, de la frecuencia, no lineal, etc.
- Selección de las características, donde se obtiene un subconjunto óptimo de características relacionadas con las fases del sueño.
- Clasificación, donde se obtienen los resultados y se distinguen las diferentes clases o estados que se quieren identificar.

Para conseguir las características deseadas se realizan diferentes tipos de análisis: temporal, frecuencial, temporal-frecuencial y no lineal [10]:

- En el análisis temporal se trabajan con parámetros estadísticos y es el más sencillo [10].
- En el análisis frecuencial se estima la potencia espectral en las diferentes bandas de frecuencia de la señal, con herramientas como la *Fast Fourier Transform* (FFT) [10].
- En la combinación de los dos métodos es común el uso de *wavelets* [10].
- En el análisis no lineal los métodos se pueden agrupar en dos grandes grupos, en primer lugar, los basados en la entropía y complejidad de la señal, que muestran los niveles de impureza e irregularidad y engloban, entre otras, la entropía de Reyi y la de Tallis. En segundo lugar, los métodos basados en las características fractales, que muestran el comportamiento aleatorio de la señal y buscan similitudes, en este grupo se encuentran los exponentes de Lypanov y de Hurst [10].

Tras la extracción de características, algunos estudios llevan a cabo una fase de selección de características para obtener un subconjunto de variables que proporcionen información relevante y complementaria relacionada con las fases del sueño [10]. Una vez obtenida la señal y seleccionadas las características se requiere de un método para su clasificación. Los métodos habitualmente utilizados son los siguientes:

- *K-means*: Método de clasificación que aplica *machine learning* para dividir el conjunto de observaciones en varios grupos o *clusters*. El número de grupos viene dado por el parámetro  $k$  [26], [27].
- Análisis discriminante lineal (ADL): Modelo de clasificación que trabaja con regiones de decisión, cuyas fronteras se definen mediante el uso de funciones lineales, y que busca minimizar la varianza entre los elementos pertenecientes a cada una de las regiones [28].
- *Support Vector Machine* (SVM): Algoritmo de aprendizaje supervisado que aplica *machine learning* para encontrar vectores de soporte que definan fronteras que separen, de manera óptima, diferentes conjuntos de datos [28].
- *Ensemble Learning*: Métodos que generan, a partir del conjunto de entrenamiento, los diferentes modelos de clasificación cuyas decisiones luego se combinan para dar lugar a la clasificación global. Dentro de estos métodos se diferencian principalmente tres subgrupos de algoritmos, *bagging*, *boosting* y *stacking*. En los métodos de *bagging* (*bootstrap aggregating*) se entrenan los modelos individuales de manera independiente (usando subgrupos de los datos de entrenamientos generados a partir de los originales replicando algunas de las muestras y eliminando otras). La idea detrás de los métodos de *boosting* es similar a la de *bagging*, pues en ambas se combinan varios modelos y se toma como decisión final la clasificación que más se repite entre los modelos individuales. Sin embargo, en *boosting*, los diferentes modelos no se entrenan por separado. Los métodos de *stacking* difieren de los anteriores en que en vez de decidir la clasificación final por un proceso de 'votación' equitativo entre las clasificaciones de los modelos individuales, se introduce una etapa extra con un algoritmo que se encarga de combinar las clasificaciones individuales y dar un resultado global [26].

### 2.2.3. MÉTODOS DE DEEP-LEARNING

Los métodos de *Deep Learning*, usando redes neuronales, han sido los que mejores resultados han obtenido normalmente. Concretamente, las técnicas de *deep learning* más utilizadas en el procesamiento de señales biomédicas son las siguientes:

- Redes neuronales artificiales (ANN): Modelos de procesamiento de información inspirados en las interacciones neuronales. Consisten en varias capas interconectadas con unidades matemáticas que imitan el comportamiento de las neuronas del cuerpo humano, formando una red, que combinan sus resultados con ciertos pesos para ir pasando la información de capa a capa [29].
- Redes neuronales convolucionales (CNNs): La CNN es un subtipo de la ANN inicialmente diseñada para el procesamiento de imágenes [30]. Las ANN convencionales son susceptibles a desviaciones en los datos que pueden afectar negativamente en el rendimiento predictivo del modelo. Las CNN superan estas limitaciones mediante una arquitectura multicapa que incorpora: (i) convoluciones de los datos de entrada con filtros con pesos compartidos que permiten extraer de manera automática, con un gran nivel de abstracción, características en distintas zonas de los datos de entrada; (ii) operaciones de reducción de dimensionalidad (pooling), que permiten retener la información más significativa y reducir la complejidad de la red.
- Redes neuronales recurrentes (RNNs). La RNN se adapta mejor al procesamiento de secuencias de datos, pues se tienen en cuenta las dependencias con datos anteriores a la hora de dar un resultado. Las RNN fueron originalmente diseñadas para el procesamiento automático de lenguaje, pero también han demostrado su utilidad en el análisis de series temporales [31].

## 2.3. COMPARACIÓN Y ELECCIÓN DEL MÉTODO A IMPLEMENTAR

Como se ha mencionado anteriormente, el primer paso del trabajo ha sido realizar una revisión del estado del arte de métodos de detección automática de las fases del sueño mediante *deep learning*, que es la metodología más novedosa y que ha proporcionado los mejores resultados, según he podido comprobar al revisar varios estudios. Posteriormente se eligió algún estudio que tratara las señales que vamos a utilizar en el trabajo como base para la realización de las posteriores etapas del trabajo, sirviendo de comparación. Para ello se prestó atención a los parámetros de precisión y *Cohen's Kappa* (Kappa).

La precisión mide el porcentaje del número de épocas correctamente clasificadas respecto al número total de épocas. El coeficiente Kappa mide la concordancia entre la clasificación hecha por el algoritmo y la misma clasificación realizada por un experto de manera visual [32]. Se representan en tanto por ciento y tanto por uno, respectivamente, con valores más altos implicando una mejor clasificación.

En el capítulo 6 se presentan los documentos que se han analizado en una tabla. Podemos diferenciar los estudios en diferentes grupos:

- En primer lugar, hay bastantes estudios que han utilizado la señal de EEG siguiendo estructuras diferentes. A destacar el estudio de Ebrahim [33] que utiliza una estructura CNN+TCNN (CNN temporal) consiguiendo resultados de precisión de 82,46% y un *Kappa* de 0,76 y el estudio de Sundar [34] que utiliza una estructura CNN+RNN consiguiendo resultados de precisión de 90,7%.
- En segundo lugar, hay estudios que han utilizado la señal de ECG o HRV extraída del ECG con peores resultados como el estudio de Werth [35] utilizando una CNN+RNN consiguiendo un *Kappa* de 0,43. La ventaja de ECG frente al EEG es que es su método de obtención es menos complejo y costoso.
- Por último, varios estudios recientemente han utilizado señales de pulsioximetría, siendo el estudio a destacar y que nos ha servido realmente para la realización del trabajo el estudio de Henri Korkalainen [36], utilizando la señal de PPG y dando unos valores de precisión y kappa que hemos conseguido mejorar con la clasificación de las 5 fases del sueño con una precisión de 68,79% y un *Kappa* de 0,587. Además, se ha seguido un método similar al utilizado en este trabajo, el uso de una CNN combinado con una RNN, que se ha comparado con una CNN. La principal diferencia respecto a este estudio y este trabajo es que en este trabajo se ha utilizado una base de datos de niños, mientras que en el otro estudio era de adultos. Además, en este TFG no solo se utiliza la señal de PPG, sino también las señales de PR y la de SpO<sub>2</sub>, que tienen la ventaja de ser adquiridas y almacenadas por pulsioxímetros portátiles para un posterior posprocesado. Además del estudio anterior, también se ha utilizado como base un trabajo previo realizado por el grupo de investigación de mis tutores de TFG [37], que utilizando una CNN ha conseguido unos valores de precisión de 78,2% y un *Kappa* de 0,57, clasificando 3 fases del sueño (W, REM y NREM).



# Capítulo 3: SUJETOS Y SEÑALES

## 3.1. BASE DE DATOS

En este capítulo se presenta la base de datos utilizada a lo largo del trabajo y se especifican las diferentes etapas de preprocesado necesarias. Finalmente, se explica el criterio elegido para su división en entrenamiento, validación y test.

### 3.1.1. POBLACIÓN BAJO ESTUDIO

La base de datos utilizada en el trabajo, CHAT (*Childhood Adenotonsillectomy Trial*), se diseñó con el objetivo de determinar diferencias en la forma de tratar la apnea del sueño obstructiva en niños mediante una operación de adenotonsilectomía (*early Adenotonsillctomy*, eAT) o sin ningún tipo de cirugía (*watchful waiting with supportive care*, WWSC) [34]. Los niños participantes en el estudio tenían entre 5 y 10 años de edad y los padres (o tutores legales) dieron el consentimiento para su participación en el estudio. A los niños se les realizó el estudio polisomnográfico y, de acuerdo a las reglas AASM 2007, se anotaron los eventos cardiorrespiratorios y las etapas del sueño [11].

Para este estudio se seleccionaron 453 niños de los 1447 totales, correspondientes a los sujetos del grupo de baseline y que fueron aleatorizados para eAT o WWSC. Para realizar este trabajo, se volvió a hacer un proceso de descarte de sujetos con señales de PR/ SpO<sub>2</sub> y/o PPG ruidosas, reduciendo el número de sujetos a 429 para el análisis de las señales de PR+ SpO<sub>2</sub>, cuyas características demográficas se pueden ver en la tabla 3.1 [19], y 366 sujetos para el análisis de las señales de PPG+ SpO<sub>2</sub>, como se ve en la tabla 3.2 [19]. Para estos sujetos se dispone, además de las señales de pulsioximetría, de las etiquetas con las fases del sueño correctamente diferenciadas.

Tabla 3.1: Características de la base de datos para las señales de PR+SpO<sub>2</sub>.

Características base de datos PR+ SpO <sub>2</sub>	
Número de sujetos (n)	429
Media de edad (años)	6
Número de niños (n)	208 (48,5%)
BMI (kg/m <sup>2</sup> )	17,2
AHI (e/h)	4,8

Tabla 3.2: Características de la base de datos para las señales de PPG+SpO<sub>2</sub>.

Características base de datos PPG+ SpO <sub>2</sub>	
Número de sujetos (n)	366
Media de edad (años)	6
Número de niños (n)	178 (49%)
BMI (kg/m <sup>2</sup> )	17,2
AHÍ (e/h)	4,8

### 3.1.2. CARACTERÍSTICAS DE LAS SEÑALES DE PULSIOXIMETRÍA

Para iniciar el proceso de preprocesado se extrajeron las señales de pulsioximetría del registro polisomnográfico en EDF a un formato .MAT, que es más manejable por el software de procesado. Posteriormente, se realiza el preprocesado que varía ligeramente en función de la combinación de señales utilizada.

- PR, SpO<sub>2</sub>, y combinación de ambas (PR+ SpO<sub>2</sub>): Se realizó la extracción de las señales por separado y de las etiquetas de las fases del sueño. Estas señales se remuestran a una frecuencia de 1Hz para homogeneizar las señales, siendo 1Hz una frecuencia de muestreo habitual en pulsioxímetros. A continuación, se ha realizado una normalización por sujeto de cada señal mediante *z-score*. Finalmente, las señales se dividen en épocas de 30 segundos asignado a cada época su fase de sueño correspondiente, para su posterior procesado. Como se mencionó anteriormente las fases N3 Y N4 se han unido. Las señales y las etiquetas son extraídas y colocadas en dos matrices bidimensionales (número de épocas x número de muestras por época), una para cada señal. Por último, las matrices bidimensionales se juntan en una tridimensional para juntar las dos señales.
- PPG y PPG con SpO<sub>2</sub> (PPG+SpO<sub>2</sub>): Se realizó un proceso similar con la diferencia de que se utilizó una frecuencia de muestreo 32 Hz para ambas señales, que ha demostrado ser adecuada para el análisis de la señal de PPG en estudios previos [38].

## 3.2. DIVISIÓN EN ENTRENAMIENTO, VALIDACIÓN Y TEST

Las bases de datos utilizadas para conseguir modelos de *deep learning* se dividen en tres grupos: datos de entrenamiento, datos de validación y datos de test. Cada uno de los sujetos pertenece únicamente a un grupo y, por lo tanto, es independiente del resto de grupos:

- Entrenamiento: El grupo de entrenamiento se utiliza para el proceso de entrenamiento de una red neuronal, cuyo fin es conseguir un modelo de clasificación de la mejor precisión posible. Este grupo es el más grande porque necesita de más datos para realizar la optimización de los pesos del modelo de *deep learning* correspondiente
- Validación: El grupo de validación sirve para controlar la evolución de los modelos entrenados y es fundamental para elegir los valores óptimos de los hiperparámetros de la red, de los cuáles se hablará más tarde. Este grupo ha sido clave ya que las diferentes pruebas con los diferentes modelos utilizados se han realizado y decidido en base a los resultados con este grupo.
- Test: Una vez decidido el modelo que mejor se ha comportado con el grupo de validación se procede a probar el modelo con el último grupo, el de test, cuyos resultados son los finales del trabajo. Este grupo proporciona la medida más fiable en un grupo independiente.

En este estudio se ha realizado un procedimiento de división muy sencillo. Se ha decidido usar un 60% de los sujetos para el grupo de entrenamiento, un 20% para el de validación y un 20% para el de test, siguiendo un poco la división realizada en un trabajo previo [37]. La selección de los sujetos ha sido realizada por orden de numeración en la base de datos. Con estos porcentajes de división conseguimos que los registros usados en el procedimiento de entrenamiento sean suficientemente grandes para conseguir unos modelos de clasificación con buenos resultados.

En las tablas 3.3 y 3.4 podemos observar como el procedimiento de selección varía ligeramente en función del tipo de señal (PR+SpO<sub>2</sub> o PPG+SpO<sub>2</sub>) debido al número de señales (sujetos) disponibles.

*Tabla 3.3: División de señales para entrenamiento, validación y test de las combinaciones PR, SpO<sub>2</sub> y de PR+SpO<sub>2</sub>*

PR+SpO <sub>2</sub>	
TOTALES	453
VÁLIDAS	429
ENTRENAMIENTO	258
VALIDACIÓN	85
TEST	86

*Tabla 3.4: División de señales para entrenamiento, validación y test de las combinaciones PPG y PPG+SpO<sub>2</sub>.*

PPG+SpO <sub>2</sub>	
TOTALES	392
VÁLIDAS	366
ENTRENAMIENTO	219
VALIDACIÓN	73
TEST	74



# Capítulo 4: METODOLOGÍA

Una red neuronal es un modelo computacional que emula el modo en el que el cerebro humano procesa la información. Funciona mediante un número elevado de unidades de procesamiento interconectadas que se organizan en capas. Las partes principales son: capa de entrada, capas ocultas y capa de salida. La red aprende examinando registros individuales y generando predicciones para cada registro. Hay parámetros que permiten realizar ajustes para corregir las predicciones [17]. El proceso de aprendizaje se conoce como entrenamiento. En la figura 4.1 [39] podemos observar como este modelo de procesamiento es característico del *deep learning*.

## 4.1. REDES NEURONALES CONVOLUCIONALES

Una red neuronal convolucional (CNN) es un tipo de red neuronal profunda [36]. Está constituida por una gran cantidad de unidades de procesamiento, neuronas, asociadas a parámetros de corrección (*bias*) y de pesos (filtros o *kernel*). El proceso de entrenamiento se realiza para obtener el valor óptimo de estos parámetros. Cada neurona recibe varias entradas, se realiza una suma ponderada en función de los pesos, se pasa por una función de activación y se genera una salida. Podemos ver un ejemplo en la figura 4.2 [17]. El uso de las CNN destaca en el reconocimiento de imágenes [40].

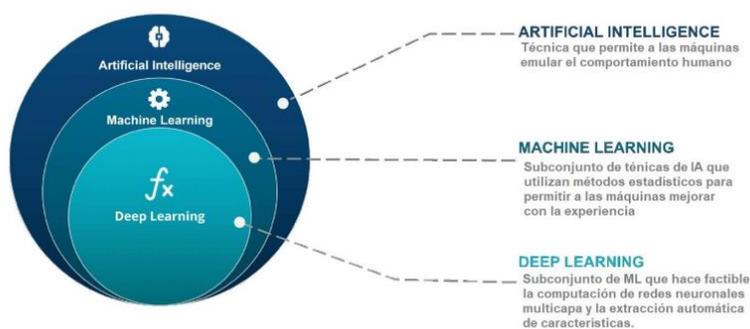


Figura 4.1: Diagrama de Inteligencia Artificial, Machine Learning y Deep Learning.

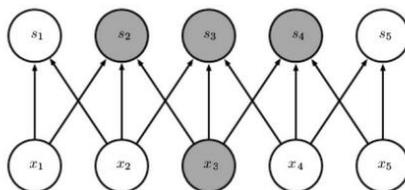


Figura 4.2: Ejemplo de conexión entre neuronas en una CNN.

Las ideas fundamentales bajo este tipo de redes son la *sparse connectivity* y el *weight sharing* [17]. La primera idea implica que cada neurona de la capa está conectada a tantas neuronas como indique el tamaño del filtro, permitiendo que la red describa de manera eficiente las diferentes interacciones complejas, creando pequeños bloques de menor tamaño. La segunda idea implica que las CNNs hagan uso compartido de parámetros al realizar las convoluciones, aprendiendo un único conjunto para todas y reduciendo el almacenamiento necesario.

La arquitectura de una capa de una CNN suele estar formada por tres etapas: la etapa de convolución, la etapa de reducción de dimensionalidad (*pooling*) y la etapa de interconexión completa (*fully connected*) [17].

- Etapa convolucional: Es la encargada de la extracción automática de las características. Obtiene los conocidos como *feature maps* (mapas de características) gracias a las convoluciones realizadas entre los datos de entrada a la capa y el *kernel*. Las CNN usan convoluciones discretas. Dentro de esta operación se lleva a cabo un proceso de *padding* o relleno para igualar el número de elementos del vector mediante ceros. Este *padding* puede ser *full*, *same* o *valid* [41]. El más utilizado es el *same* ya que hace que el vector de salida tenga las mismas dimensiones que el de entrada. Además, se puede realizar un proceso de submuestreo para reducir la dimensión de la matriz de salida. Esto es equivalente a aplicar un *stride* (desplazamiento de los filtros al ir convolucionando los datos de entrada). Un mayor *stride* implica menos tiempo y menos convoluciones. Tras la convolución, la amplitud de los *feature maps* se suelen normalizar mediante una operación conocida como *batch normalization*. A continuación, una función de activación, se encarga de umbralizar las características extraídas en la capa anterior. Se puede realizar mediante diferentes funciones, lineales o no lineales. La más utilizada es la función de activación *Rectified Linear Unit* (ReLU) [17], ya que no cambia las dimensiones, solo realiza un mapeo uno por uno de los valores de activación de cada elemento de la capa convolucional.
- Etapa de *pooling*: En esta etapa se busca que las salidas conseguidas tras cada capa de convolución no se vean afectada por pequeñas variaciones en las entradas. Esto se consigue sustituyendo el valor de cada unidad de salida de la capa convolucional por otro que tenga también en cuenta el valor de las salidas más próximas. En la figura 4.3 podemos observar los diferentes tipos de *pooling* [42]. Dependiendo del tipo de *pooling* esto se realiza de una manera u otra. Por ejemplo, en *max pooling* cada unidad toma el valor máximo entre todas las que la rodean, mientras que en *average pooling* se realiza la media.

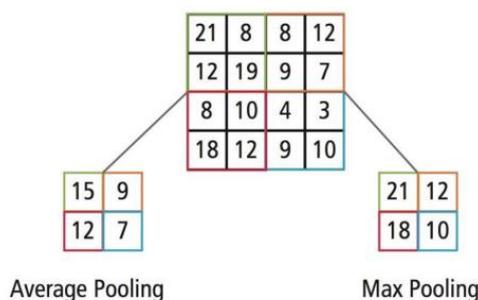


Figura 4.3: Diferencias entre Average Pooling y Max Pooling.

- Etapa *fully connected*: Todos los nodos de la capa anterior están conectados a todos los nodos de esta etapa. Es la capa(s) que va al final de la red y se encarga de implementar la clasificación empleando las características extraídas en las capas de “convolución + *pooling*” anteriores. El número total de nodos de salida en la última capa *fully connected* determina el número de clases de salida de la red.

Existen dos etapas diferenciadas a la hora de entrenar una CNN La primera es la etapa en la que se introduce la señal a clasificar en la red y ésta, tras realizar una serie de multiplicaciones o convoluciones, produce una salida y la segunda es la etapa de propagación hacia atrás (*backpropagation*), donde los pesos de la red se actualizan para que el error de clasificación cometido vaya decreciendo al avanzar en el entrenamiento [17].

#### 4.1.1. ARQUITECTURA CNN APLICADA

La programación de las redes neuronales se ha realizado mediante la librería *Keras* de Python. Se ha seguido el modelo secuencial de *Keras*, que permite construir la red neuronal capa a capa. Con el objetivo de comparar el rendimiento de las distintas señales de pulsioximetría, se han probado las siguientes configuraciones de señales de entrada de la CNN: (i) señal PR; (ii) señal SpO<sub>2</sub>; (iii) señales PR y SpO<sub>2</sub> (PR+SpO<sub>2</sub>); (iv) señal PPG; (v) señales PPG y SpO<sub>2</sub> (PPG+SpO<sub>2</sub>). Para cada una de estas configuraciones, la sección de entrada de la red es la época de 30-s a clasificar junto con x épocas anteriores y z épocas posteriores, siendo ese número de épocas posteriores y anteriores ha sido obtenido en la optimización de hiperparámetros (ver sección 5.1.1). Hay que tener en cuenta que, en el caso de configuraciones con 1 señal de entrada, se realizaron convoluciones en 1 dimensión (1D), mientras que en las configuraciones con 2 señales de entrada se utilizaron funciones de convolución en 2 dimensiones (2D). En las 5 configuraciones, el modelo CNN final consiste en las siguientes capas:

- A. Capa inicial de *batch normalization*, para normalizar los datos de entrada.
- B. Dos capas convolucionales 1D/2D (según el número de señales de entrada) con 16 filtros, un *kernel* de 7, *padding same*, *batch normalization*, activación *ReLU* y *dropout* de 0.1.
- C. Dos capas convolucionales 1D/2D de 32 filtros, un *kernel* de 5, *padding same*, *batch normalization*, activación *ReLU* y *dropout* de 0.1.
- D. Dos capas convolucionales 1D/2D de un tamaño de filtro de 64 un *kernel* de 3, *padding same*, *batch normalization*, activación *ReLU* y *dropout*.
- E. Una etapa *maxpooling* de 2.
- F. Una etapa *flatten*, para transformar los *feature maps* en datos 1D.
- G. Una etapa *dense* con función de activación *softmax* y 5 clases (el número de fases del sueño a clasificar).

Cada etapa convolucional consiste en una convolución, un *batch normalization*, una activación y un *dropout*.

Para la señal de PPG+SpO<sub>2</sub> se incluye detrás de cada etapa convolucional un *maxpooling*, ya que al ser mayor la frecuencia de muestreo (32 en PPG y PPG+SpO<sub>2</sub> vs. 1 en PR, SpO<sub>2</sub> y PR+SpO<sub>2</sub>) hay que reducir en mayor medida la dimensionalidad durante la red.

Ahora se procede a la explicación de cada capa [17]:

- Capa convolucional 1D/2D, que se encarga de extraer los *feature maps*. En la figura 4.4 se puede ver un ejemplo de convolución 2D. En este TFG se han utilizado tres tipos de bloques convolucionales (ver B, C y D de esta sección), cada uno de ellos con capas convolucionales con distinto número de filtros y *kernel*. Se han utilizado 2 capas convolucionales de los bloques B, C y D, valor obtenido en la optimización de hiperparámetros (ver sección 5.1.1). En cada capa convolucional 1D/2D deben especificarse ciertos parámetros:
  - *Filters*: número de filtros de salida en la convolución. En este TFG se ha ido aumentando el número de filtros a medida que aumentaba la profundidad de la CNN, un enfoque habitual en estas redes.
  - *Kernel\_size*: longitud de la ventana deslizante utilizada para realizar la convolución. En este TFG se ha disminuido el tamaño de filtro a medida que aumentaba la profundidad de la CNN, un enfoque habitual en estas redes.
  - *Strides*: número de espacios que debe moverse el filtro o *kernel*.
  - *Padding*: tipo de *padding*, como se ha explicado anteriormente.
- La etapa de *batch normalization* normaliza las salidas de cada capa para cada *mini-batch* del entrenamiento.
- La activación especifica la función de activación que se desea utilizar después de la normalización. En las etapas convolucionales se ha optado por la función ReLU, la más habitual en CNNs.
- El *dropout* es una técnica que se aplica a algunas salidas de la capa de red que elimina de forma aleatoria y periódica algunas de las conexiones entre neuronas, mediante un factor entre 0-1. Podemos observar un ejemplo de su funcionamiento en la figura 4.5 [43]. En este TFG, el valor de la probabilidad de dropout,  $p_{drop\_cnn}$ , ha sido uno de los hiperparámetros de la arquitectura CNN a optimizar (ver sección 5.1.1).
- La etapa de *maxpooling* que, como ya se ha explicado anteriormente, sirve para reducir la dimensionalidad y mantener las características más relevantes, minimizando así el sobreentrenamiento.
- La capa *flatten* adapta las dimensiones de una capa a las dimensiones de la siguiente.
- La capa dense es una capa de salida en la que se indica el número de clases a identificar y una función de activación. En este TFG se ha utilizado la función *softmax* con 5 elementos (clases), que convierte los valores de entrada en probabilidades de salida cuya suma es uno, permitiendo así obtener a la salida la probabilidad de cada fase del sueño (W, N1, N2, N3 y REM) para los datos de entrada.

El entrenamiento y optimización de los modelos CNN fue realizado utilizando la siguiente configuración:

- Un tamaño de batch (*batch\_size*: número de segmentos de entrada a la red en cada iteración) de 64, durante 200 épocas [17].
- El algoritmo *Adaptive Moment Estimation* (Adam) con una tasa de aprendizaje inicial (*learning rate*) de 0.005 para la optimización de los pesos y biases de la CNN.
- *Categorical cross-entropy* (clasificación múltiple) entre las fases del sueño de las épocas del batch y las probabilidades de salida de la CNN como función de pérdida a minimizar durante el entrenamiento [17].
- Funciones *ReduceLROnPlateau* y *EarlyStopping* de Keras, que permiten reducir la *learnign rate* por un factor de 2 cuando la precisión o la función de pérdida en el conjunto de validación no mejora en 10 épocas (*ReduceLROnPlateau*) y parar el entrenamiento después de 30 épocas sin mejoras en el grupo de validación (*EarlyStopping*), respectivamente [17].

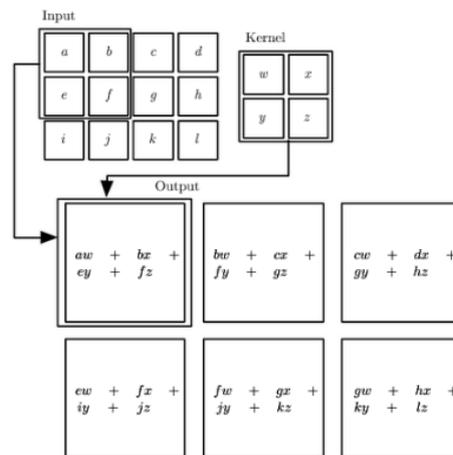


Figura 4.4: Ejemplo de convolución 2-D.

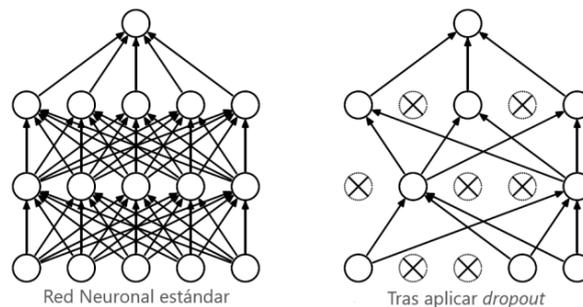


Figura 4.5: Ejemplo de una red neuronal sin dropout y con dropout.

## 4.2. REDES NEURONALES RECURRENTE

Las redes neuronales recurrentes (RNNs) son una clase de redes de *deep learning*, que incluyen conexiones hacia adelante y hacia atrás, es decir, retroalimentaciones entre las neuronas dentro de las capas. Tratan los datos de forma secuencial y recuerdan las salidas anteriores y/o posteriores como entrada. Su uso destaca en el reconocimiento de voz. En la figura 4.6 podemos observar un ejemplo de la recursividad en las RNNs [31].

En cada instante de tiempo, cada neurona recibe la entrada de la capa anterior, así como su propia salida del instante de tiempo anterior para generar su salida. En una capa de estas neuronas, cada neurona recibe dos entradas. Cada neurona tiene una cierta memoria, *memory cell*, útil para tratar datos secuenciales. Además, se realiza un proceso de *backpropagation through time* (BPTT), en el que la información se pasa de un instante a otro [31].

El tipo más común de red RNN hoy en día es la *Long-Short Term Memory* (LSTM), una extensión de las redes neuronales recurrentes, que básicamente amplían su memoria para aprender de experiencias importantes que han pasado hace mucho tiempo. Las LSTM permiten a las RNN recordar sus entradas durante un largo período de tiempo. *Keras* ofrece también otras implementaciones de RNN como es la *Gated Recurrent Unit* (GRU). Las capas GRU aparecieron en el 2014, y usan el mismo principio que LSTM, pero están simplificadas de manera que su rendimiento está a la par con LSTM, pero computacionalmente son más eficientes [31].

### 4.2.1 ARQUITECTURA CNN+RNN APLICADA

La arquitectura aplicada consiste en la combinación de una CNN y una RNN. La sección de entrada consiste en una secuencia de  $n_{seq}$  épocas consecutivas de 30 s de las señales de entrada. La arquitectura propuesta procesa de manera independiente cada época con una capa de distribución temporal (*Time Distributed*) que contiene la red CNN descrita en el apartado 4.1.1, exceptuando la capa *dense* y la función de activación *softmax*. La parte CNN se encarga de aprender las características de las fases del sueño que contiene cada época. La capa *Time Distributed CNN* es procesada mediante una RNN, para así modelar la distribución de las fases del sueño durante la secuencia. En primer lugar, la RNN contiene una capa de *dropout* gaussiano con una probabilidad de 0.3 que se encarga de reducir el *overfitting*. A continuación, una capa bidireccional de tipo GRU con  $n_{gru\_units}$  unidades se encarga de modelar la dependencia temporal de la secuencia de épocas de las señales de pulsioximetría, decidiendo qué información debe ser recordada de la red. El tamaño de secuencia ( $n_{seq}$ ) y el número de unidades GRU ( $n_{gru\_units}$ ) han sido los hiperparámetros de la arquitectura CNN+RNN a optimizar (ver sección 5.1.2) [36].

El entrenamiento y optimización del modelo CNN+RNN se ha realizado utilizando la misma configuración que con el modelo anterior CNN.

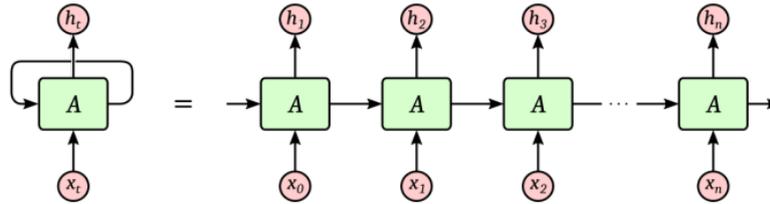


Figura 4.6: Ejemplo de recursividad en una RNN.

### 4.3. ANÁLISIS ESTADÍSTICO

En este apartado se va a describir los parámetros estadísticos utilizados para evaluar los resultados obtenidos en la clasificación automática de fases del sueño, así como en la estimación de los parámetros polisomnográficos.

#### 4.3.1. MÉTRICAS DE RENDIMIENTO DE LA CLASIFICACIÓN DE FASES DEL SUEÑO

Los resultados obtenidos por las arquitecturas CNN y CNN+RNN entrenadas con las distintas configuraciones de señales de entrada para la detección automática de las fases del sueño han sido analizados mediante los siguientes parámetros estadísticos:

- Matriz de confusión: Herramienta que permite la visualización del desempeño de un algoritmo, comparando las fases del sueño de la PSG con las predichas por los métodos de *deep learning*.
- Precisión (*accuracy*): Porcentaje de fases del sueño correctamente predichas por el algoritmo en cuestión [44].
- Coeficiente Kappa (Cohen's kappa): Es una medida estadística que se ajusta al efecto del azar en la proporción de concordancia observada para elementos cualitativos (1). Evalúa la concordancia o reproducibilidad. Valor entre 0-1, siendo el 1 el mejor valor [45].

$$k = \frac{Po - Pe}{1 - Pe} \quad (1)$$

Siendo  $P_o$  la proporción de acuerdos observados y  $P_e$  la proporción de acuerdos esperados en la hipótesis de independencia entre los observadores, es decir, de acuerdos por azar.

- MF1: Media armónica del valor predictivo positivo (*precision*) y la sensibilidad (*recall*), entre los valores 0-1 siendo la mejor puntuación 1 [44].

$$F1 = 2 * \frac{precision * recall}{precision + recall} \quad (2)$$

- *Per-class F1 score*: MF1 para cada clase individual (W, N1, N2, N3 y REM) [46].

#### 4.3.2. MÉTRICAS DE RENDIMIENTO DE LA ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS

A partir de la salida de las arquitecturas CNN y CNN+RNN, se ha estimado el porcentaje de tiempo en cada fase del sueño, así como el tiempo total del sueño (TST). Estos valores estimados se han comparado con los de la PSG mediante las siguientes métricas:

- *Scatter-plot*: Gráfico de dispersión. Muestra la intensidad de la relación entre dos variables numéricas. En este trabajo se han realizado para las diferentes fases del sueño (N1, N2, N3 y REM) y el *Total Sleep Time* (TST).
- Coeficiente de correlación intraclase (*Intraclass correlation coefficient*, ICC): Permite medir la concordancia entre dos o más valoraciones cuantitativas. En este trabajo se utiliza para evaluar la concordancia entre los parámetros polisomnográficos estimados y los de la PSG. [47]
- *p*-valor: El *p*-valor nos muestra la probabilidad de haber obtenido el resultado que hemos obtenido suponiendo que la hipótesis nula es cierta. En este trabajo se utiliza el de Wilcoxon para evaluar si hay diferencias estadísticas entre los parámetros polisomnográficos estimados y los de la PSG [37].
- Error medio: Media del error entre el resultado real y el predicho. Se calcula realizando la media de la resta de los valores predichos menos los valores reales.
- Coeficiente de correlación de Spearman: Medida no paramétrica de la correlación de rango. Mide la fuerza y dirección de la asociación entre dos variables. Puede tomar un valor entre -1 y +1 donde el primero significa una asociación perfecta negativa y el último una asociación perfecta positiva. Un valor de 0 muestra que no hay asociación. [48].

# Capítulo 5: RESULTADOS

En este capítulo se van a mostrar los resultados obtenidos en este TFG. Se va a realizar una evaluación usando el grupo de sujetos de test para los diferentes modelos entrenados y optimizados con los grupos de entrenamiento y validación, respectivamente. En primer lugar, se muestra la optimización de hiperparámetros de las arquitecturas CNN y CNN+RNN. Posteriormente, se muestra la clasificación de fases del sueño de ambas arquitecturas para las 5 combinaciones de señales diferentes: PR, SpO<sub>2</sub>, PR+SpO<sub>2</sub> (1Hz), PPG y PPG+SpO<sub>2</sub> (32 Hz). Por último, se muestran los resultados obtenidos en la estimación de parámetros polisomnográficos.

## 5.1. OPTIMIZACIÓN DE HIPERPARÁMETROS

Los conjuntos de entrenamiento y validación se han utilizado para optimizar las arquitecturas CNN y CNN+RNN. El conjunto de entrenamiento fue utilizado para entrenar las arquitecturas CNN y CNN+RNN con los distintos valores de los hiperparámetros, mientras que en el conjunto de validación se evaluó el rendimiento de estas arquitecturas para obtener los valores óptimos de los hiperparámetros que maximizan la precisión.

### 5.1.1. OPTIMIZACIÓN DE LA ARQUITECTURA CNN

A partir de la arquitectura CNN descrita en la sección 4.1.1, se procedió a realizar pruebas con diferentes valores para los siguientes hiperparámetros: (i) número de bloques convolucionales ( $n\_bloques\_cnn$ ); (ii) número de épocas anteriores (x) y posteriores (z) de las señales de entrada a solapar: y (iii) *dropout* ( $p\_drop\_cnn$ ). Los resultados en el conjunto de validación de las pruebas que se realizaron con la configuración de señales de entrada PR+SpO<sub>2</sub> se muestran en las tablas 5.1 y 5.2. En cuanto al número de épocas a solapar (ver Tabla 5.1), los mejores resultados se obtuvieron con la combinación de 5 épocas anteriores y 4 posteriores. Con este valor se procedió a realizar pruebas con los diferentes valores de  $p\_drop\_cnn$  y  $n\_bloques\_cnn$ , siendo el *dropout* entre los valores 0.1, 0.2 y 0.3 y  $n\_bloques\_cnn$  de cada tamaño de filtro 1, 2, y 3. Los mejores resultados se obtuvieron con  $p\_drop\_cnn = 0.1$  y la estructura de  $n\_bloques\_cnn = 2$  bloques previamente seleccionada (ver Tabla 5.2).

Tabla 5.1: Resultados de precisión obtenidos con las diferentes etapas de solapamiento.

Etapas solapadas x-z	Precisión
Sin solapamiento	0,5352
3-2	0,6115
4-3	0,5841
5-4	0,6314
7-6	0,5859

Tabla 5.2: Resultados de precisión obtenidos con los diferentes valores de dropout y bloques convolucionales.

<i>p_drop_cnn</i> y <i>n_bloques_cnn</i>	Precisión
<i>p_drop_cnn</i> : 0.1, <i>n_bloques_cnn</i> : 1	0,616
<i>p_drop_cnn</i> : 0.1, <i>n_bloques_cnn</i> : 2	0,631
<i>p_drop_cnn</i> : 0.1, <i>n_bloques_cnn</i> : 3	0,624
<i>p_drop_cnn</i> : 0.2, <i>n_bloques_cnn</i> : 1	0,589
<i>p_drop_cnn</i> : 0.2, <i>n_bloques_cnn</i> : 2	0,593
<i>p_drop_cnn</i> : 0.2, <i>n_bloques_cnn</i> : 3	0,616
<i>p_drop_cnn</i> : 0.3, <i>n_bloques_cnn</i> : 1	0,582
<i>p_drop_cnn</i> : 0.3, <i>n_bloques_cnn</i> : 2	0,585
<i>p_drop_cnn</i> : 0.3, <i>n_bloques_cnn</i> : 3	0,616

### 5.1.2. OPTIMIZACIÓN DE LA ARQUITECTURA CNN+RNN

A partir de la arquitectura CNN+RNN descrita en la sección 4.2.1, se han realizado pruebas de optimización con dos hiperparámetros: (i) el tamaño de secuencia (*n\_seq*), con valores de 25, 50 y 100; y (ii) el número de unidades GRU (*gru\_units*), con valores de 16, 32 y 64. Los resultados en el conjunto de validación de las pruebas que se realizaron con la configuración de señales de entrada PR+SpO<sub>2</sub> se muestran en la tabla 5.3. Los mejores resultados se obtuvieron con los siguientes valores: *n\_seq*=100 y *n\_gru\_units*=16. Posteriormente se realizó una prueba para ver si sustituyendo la capa GRU por una LSTM se mejoraban los resultados, resultado negativo la mejora.

## 5.2. CLASIFICACIÓN DE FASES DEL SUEÑO

Una vez ajustados los hiperparámetros de las arquitecturas CNN y CNN+RNN, se procedió a evaluar el rendimiento diagnóstico de estas en el conjunto de test para las 5 configuraciones de señales: PR, SpO<sub>2</sub>, PR+SpO<sub>2</sub>, PPG y PPG+ SpO<sub>2</sub>.

Tabla 5.3: Resultados de precisión obtenidos con los diferentes tamaños de secuencias y unidades GRU.

<i>n_seq</i> y <i>n_gru_units</i>	PRECISIÓN
<i>n_seq</i> : 25, <i>n_gru_units</i> :16	0,649
<i>n_seq</i> : 25, <i>n_gru_units</i> :32	0,641
<i>n_seq</i> : 25, <i>n_gru_units</i> :64	0,631
<i>n_seq</i> : 50, <i>n_gru_units</i> :16	0,663
<i>n_seq</i> : 50, <i>n_gru_units</i> :32	0,653
<i>n_seq</i> : 50, <i>n_gru_units</i> :64	0,657
<i>n_seq</i> : 100, <i>n_gru_units</i> :16	0,669
<i>n_seq</i> : 100, <i>n_gru_units</i> :32	0,654
<i>n_seq</i> : 100, <i>n_gru_units</i> :64	0,661

## 5.2.1 CLASIFICACIÓN DE FASES DEL SUEÑO DE LA ARQUITECTURA CNN

En primer lugar, se muestran las matrices de confusión de las arquitecturas CNN en el conjunto de test en las figuras 5.1 y 5.2. A partir de estas figuras se han obtenido los valores de precisión, *Kappa*, *MF1* y *per-class F1-score* en el conjunto de test, mostrados en la tabla 5.4. Es importante destacar que la señal de PPG alcanzó los mejores resultados, superando tanto a las señales PR y SpO<sub>2</sub> individualmente, como a la combinación de ellas y a la combinación PPG+SpO<sub>2</sub>.

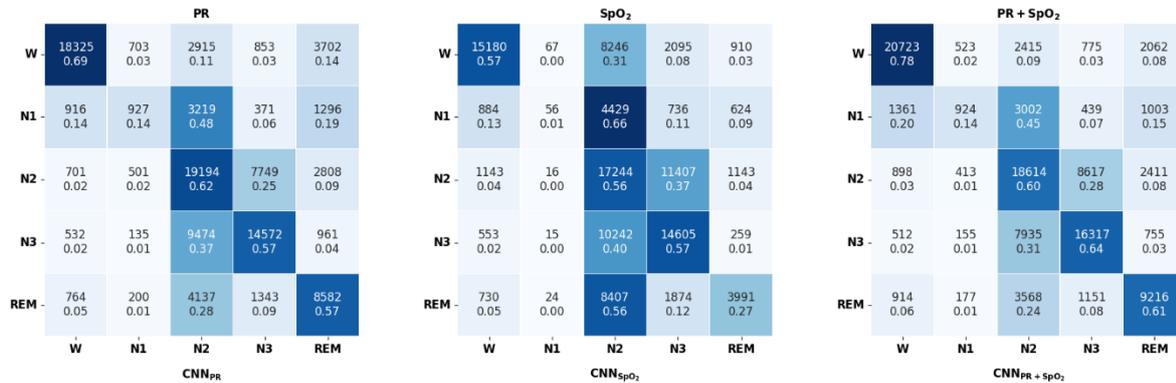


Figura 5.1: Matrices de confusión de las señales PR, SpO<sub>2</sub> y PR+SpO<sub>2</sub> obtenidas con CNN.

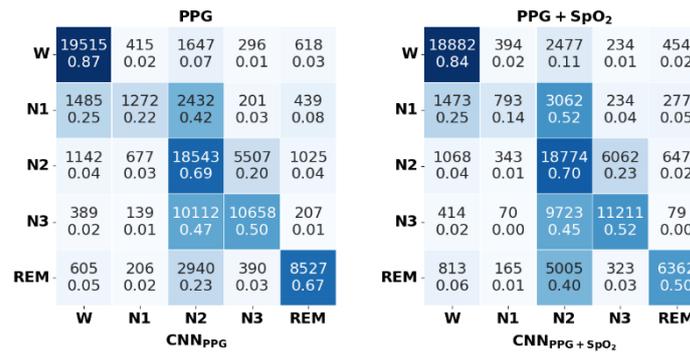


Figura 5.2: Matrices de confusión de las señales PPG y PPG+SpO<sub>2</sub> obtenidas con CNN.

Tabla 5.4: Resultados de precisión, Kappa, MF1 y per-class F1-score en test para la estructura CNN.

SEÑALES	PRECISIÓN	KAPPA	MF1	Per-class F1-score				
				W	N1	N2	N3	REM
PR+SpO <sub>2</sub>	0,6273	0,5064	0,5604	0,81	0,21	0,56	0,62	0,6
PR	0,5873	0,4541	0,5250	0,77	0,2	0,55	0,58	0,53
SpO <sub>2</sub> (1HZ)	0,4870	0,3044	0,4013	0,67	0,02	0,43	0,52	0,36
PPG+SpO <sub>2</sub>	0,6267	0,4972	0,5604	0,84	0,21	0,57	0,57	0,62
PPG	0,6546	0,5394	0,6050	0,86	0,3	0,59	0,55	0,73

## 5.2.2 CLASIFICACIÓN DE FASES DEL SUEÑO DE LA ARQUITECTURA CNN+RNN

En primer lugar, se muestran las matrices de confusión de las arquitecturas CNN+RNN en el conjunto de test en las figuras 5.3 y 5.4. A partir de estas figuras se han obtenido los valores de precisión, *Kappa*, *MF1* y *per-class F1-score* en el conjunto de test, mostrados en la tabla 5.5. Es importante destacar que todos los resultados mejoran respecto al uso único de la CNN (sección 5.2.1). Con esta arquitectura, la señal de PPG alcanzó también los mejores resultados, aunque con una mejora con respecto a las demás configuraciones de señales mucho menor que en la sección anterior.

PR						SpO <sub>2</sub>						PR + SpO <sub>2</sub>					
W	21176 0.80	717 0.03	2845 0.11	245 0.01	1515 0.06	W	19738 0.74	185 0.01	3013 0.11	1261 0.05	2301 0.09	W	21107 0.80	848 0.03	2527 0.10	487 0.02	1529 0.06
N1	1292 0.19	1032 0.15	3545 0.53	181 0.03	679 0.10	N1	1815 0.27	175 0.03	2230 0.33	615 0.09	1894 0.28	N1	1317 0.20	1051 0.16	3217 0.48	329 0.05	815 0.12
N2	893 0.03	540 0.02	23363 0.75	4008 0.13	2149 0.07	N2	3658 0.12	218 0.01	14163 0.46	7622 0.25	5292 0.17	N2	1029 0.03	512 0.02	19930 0.64	7327 0.24	2155 0.07
N3	423 0.02	104 0.00	9727 0.38	14983 0.58	437 0.02	N3	1742 0.07	43 0.00	7734 0.30	14334 0.56	1821 0.07	N3	533 0.02	118 0.00	5304 0.21	19379 0.75	340 0.01
REM	717 0.05	247 0.02	4350 0.29	307 0.02	9405 0.63	REM	2399 0.16	138 0.01	4001 0.27	1444 0.10	7044 0.47	REM	561 0.04	248 0.02	3415 0.23	859 0.06	9943 0.66
	W	N1	N2	N3	REM	W	N1	N2	N3	REM	W	N1	N2	N3	REM		
	CNN + RNN <sub>PR</sub>					CNN + RNN <sub>SpO<sub>2</sub></sub>					CNN + RNN <sub>PR + SpO<sub>2</sub></sub>						

Figura 5.3: Matrices de confusión de las señales PR, SpO<sub>2</sub> Y PR+ SpO<sub>2</sub> obtenidas con CNN+RNN.

PPG						PPG + SpO <sub>2</sub>					
W	19800 0.88	610 0.03	1137 0.05	157 0.01	787 0.03	W	18514 0.82	779 0.03	1766 0.08	312 0.01	1120 0.05
N1	1744 0.30	1364 0.23	2023 0.35	182 0.03	516 0.09	N1	1169 0.20	948 0.16	2767 0.47	326 0.06	619 0.11
N2	1983 0.07	715 0.03	17264 0.64	5436 0.20	1496 0.06	N2	1009 0.04	395 0.01	17294 0.64	6379 0.24	1817 0.07
N3	931 0.04	111 0.01	5919 0.28	14347 0.67	197 0.01	N3	222 0.01	103 0.00	5358 0.25	15073 0.70	749 0.03
REM	952 0.08	233 0.02	2428 0.19	345 0.03	8710 0.69	REM	457 0.04	159 0.01	3056 0.24	860 0.07	8136 0.64
	W	N1	N2	N3	REM	W	N1	N2	N3	REM	
	CNN + RNN <sub>PPG</sub>					CNN + RNN <sub>PPG + SpO<sub>2</sub></sub>					

Figura 5.4: Matrices de confusión de las señales PPG y PPG+SpO<sub>2</sub> obtenidas con CNN+RNN.

Tabla 5.5: Resultados de precisión, *Kappa*, *MF1* y *per-class F1-score* en test para la estructura CNN+RNN.

SEÑALES	PRECISIÓN	KAPPA	MF1	<i>Per-class F1-score</i>				
				W	N1	N2	N3	REM
PR+SpO <sub>2</sub>	0,6809	0,5778	0,6084	0,83	0,22	0,61	0,72	0,67
PR	0,6670	0,5564	0,5959	0,83	0,22	0,62	0,66	0,64
SpO <sub>2</sub> (1HZ)	0,5287	0,3776	0,4389	0,71	0,05	0,46	0,56	0,42
PPG+SpO <sub>2</sub>	0,6708	0,5639	0,6013	0,84	0,23	0,61	0,68	0,65
PPG	0,6879	0,5870	0,6306	0,83	0,31	0,62	0,68	0,71

## 5.3. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS

### 5.3.1. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS A PARTIR DE LA ARQUITECTURA CNN

Se muestran el  $p$ -valor, el ICC, el error medio y el índice de correlación obtenido para la estimación del porcentaje de tiempo en cada fase del sueño: N1 (tabla 5.6), N2 (tabla 5.7), N3 (tabla 5.8), REM (tabla 5.9) y Wake (tabla 5.10).

*Tabla 5.6: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase N1 para la estructura CNN.*

N1				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	<0,001	0,1668	-0,0434	0,252
PR	<0,001	0,1889	-0,0411	0,314
SpO <sub>2</sub>	<0,001	0,0180	-0,0629	0,208
PPG + SpO <sub>2</sub>	<0,001	0,2206	-0,0462	0,321
PPG	<0,001	0,2547	-0,0358	0,228

*Tabla 5.7: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase N2 para la estructura CNN.*

N2				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	<0,001	0,2737	0,0435	0,280
PR	<0,001	0,2920	0,0788	0,354
SpO <sub>2</sub>	<0,001	0,1224	0,1668	0,133
PPG + SpO <sub>2</sub>	<0,001	0,3004	0,1377	0,384
PPG	<0,001	0,4053	0,0989	0,470

*Tabla 5.8: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase N3 para la estructura CNN.*

N3				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	0,176	0,0261	0,0161	-0,008
PR	0,313	-0,0213	-0,0093	-0,033
SpO <sub>2</sub>	0,003	0,0981	0,0494	0,135
PPG + SpO <sub>2</sub>	0,017	0,1551	-0,0390	0,134
PPG	<0,001	0,1718	-0,0509	0,164

**Tabla 5.9: Resultados del p-valor, ICC, error medio e índice de correlación para la fase REM para la estructura CNN.**

REM				
Señales	p-valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	0,957	0,0084	0,0034	0,020
PR	0,009	0,0358	0,0221	0,086
SpO <sub>2</sub>	<0,001	0,0600	-0,0778	0,052
PPG + SpO <sub>2</sub>	<0,001	0,4200	-0,0534	0,360
PPG	0,003	0,5098	-0,0195	0,488

**Tabla 5.10: Resultados del ICC y error medio para la fase W para la estructura CNN.**

W		
Señales	ICC	Error medio
PR + SpO <sub>2</sub>	0,7931	-0,0195
PR	0,6863	-0,0505
SpO <sub>2</sub>	0,5100	-0,0755
PPG + SpO <sub>2</sub>	0,7431	8,5228e-04
PPG	0,6493	0,0073

En la fase *Wake* se obtienen los mejores resultados y es el mejor estimado. En base a ello, se muestra el gráfico de dispersión para el *Total Sleep Time* (TST), que se puede obtener a partir del porcentaje de tiempo en *Wake* y el tiempo de registro, en las figuras 5.5 y 5.6. Se observa como exista una alta concordancia entre el TST estimado por las arquitecturas CNN y el obtenido durante la PSG.

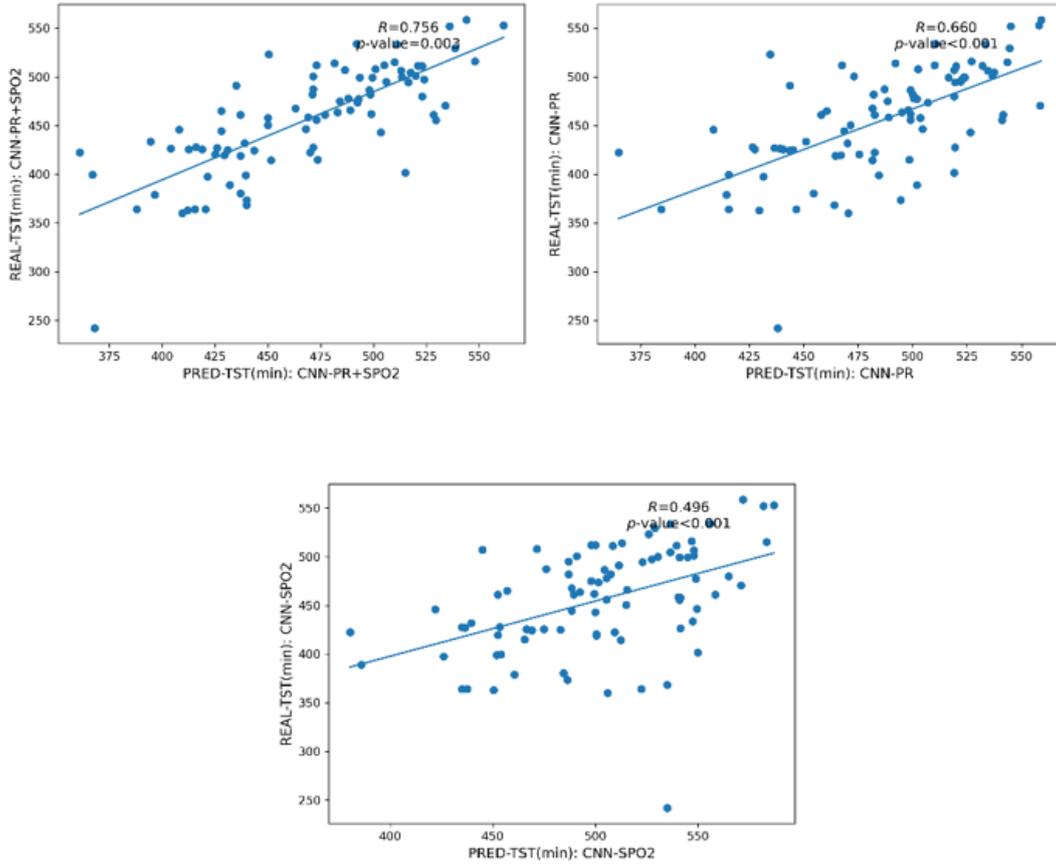


Figura 5.5: Gráfico de dispersión de las señales PR,  $SpO_2$  y PR+  $SpO_2$  para el TST obtenidos con CNN.

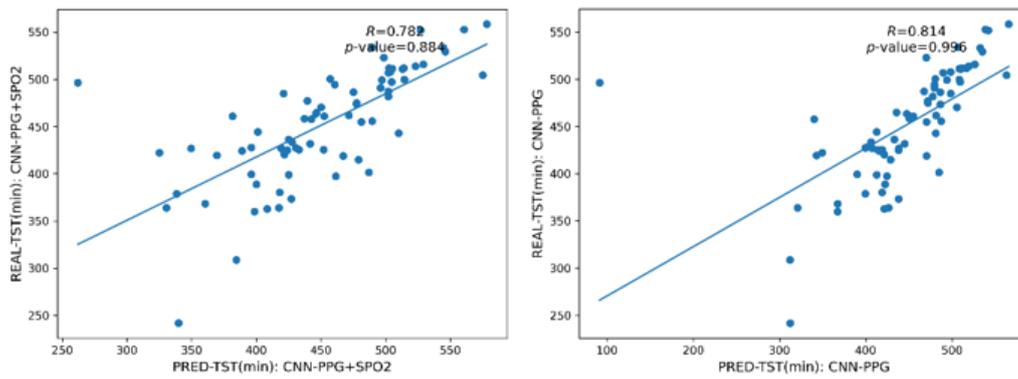


Figura 5.6: Gráfico de dispersión de las señales PPG y PPG+  $SpO_2$  para el TST obtenidos con CNN.

### 5.3.2. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS A PARTIR DE LA ARQUITECTURA CNN+RNN

Se muestran el  $p$ -valor, el ICC, el error medio y el índice de correlación obtenido para la estimación del porcentaje de tiempo en cada fase del sueño: N1 (tabla 5.11), N2 (tabla 5.12), N3 (tabla 5.13), REM (tabla 5.14) y Wake (tabla 5.15).

**Tabla 5.11: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase N1 para la estructura CNN+RNN.**

N1				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	<0,001	0,1492	-0,0375	0,285
PR	<0,001	0,2642	-0,0391	0,401
SpO <sub>2</sub>	<0,001	0,0159	-0,0574	0,123
PPG + SpO <sub>2</sub>	<0,001	0,3295	-0,0393	0,367
PPG	<0,001	0,2923	-0,0321	0,294

**Tabla 5.12: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase N2 para la estructura CNN+RNN.**

N2				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	0,001	0,3403	0,0321	0,287
PR	<0,001	0,3151	0,1226	0,326
SpO <sub>2</sub>	0,798	0,2652	0,0032	0,286
PPG + SpO <sub>2</sub>	0,081	0,2658	0,0340	0,348
PPG	0,115	0,3375	0,0184	0,379

**Tabla 5.13: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase N3 para la estructura CNN+RNN.**

N3				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	0,004	0,1915	0,0262	0,181
PR	<0,001	0,0779	-0,0558	0,062
SpO <sub>2</sub>	0,565	0,1955	-0,0043	0,297
PPG + SpO <sub>2</sub>	0,169	0,1485	0,0201	0,240
PPG	0,525	0,2000	-0,0117	0,155

**Tabla 5.14: Resultados del  $p$ -valor, ICC, error medio e índice de correlación para la fase REM para la estructura CNN+RNN.**

REM				
Señales	$p$ -valor	ICC	Error medio	Índice de correlación
PR + SpO <sub>2</sub>	0,458	0,1998	-0,0026	0,261
PR	0,243	0,1703	-0,0085	0,191
SpO <sub>2</sub>	0,161	-0,0664	0,0283	-0,085
PPG + SpO <sub>2</sub>	0,174	0,2878	-0,0026	0,434
PPG	0,131	0,4437	-0,0092	0,490

Tabla 5.15: Resultados del ICC y error medio para la fase W para la estructura CNN+RNN.

W		
Señales	ICC	Error medio
PR + SpO <sub>2</sub>	0,7557	-0,0182
PR	0,8456	-0,0193
SpO <sub>2</sub>	0,3780	0,0302
PPG + SpO <sub>2</sub>	0,7278	-0,0122
PPG	0,4475	0,0346

Como se he observado en la sección anterior, en la fase Wake se obtienen los mejores resultados y es el mejor estimado. En base a ello, se muestra el gráfico de dispersión para el *Total Sleep Time* (TST) en las figuras 5.7 y 5.8, donde también se observa una alta concordancia entre el TST estimado por las arquitecturas CNN+RNN y el obtenido durante la PSG.

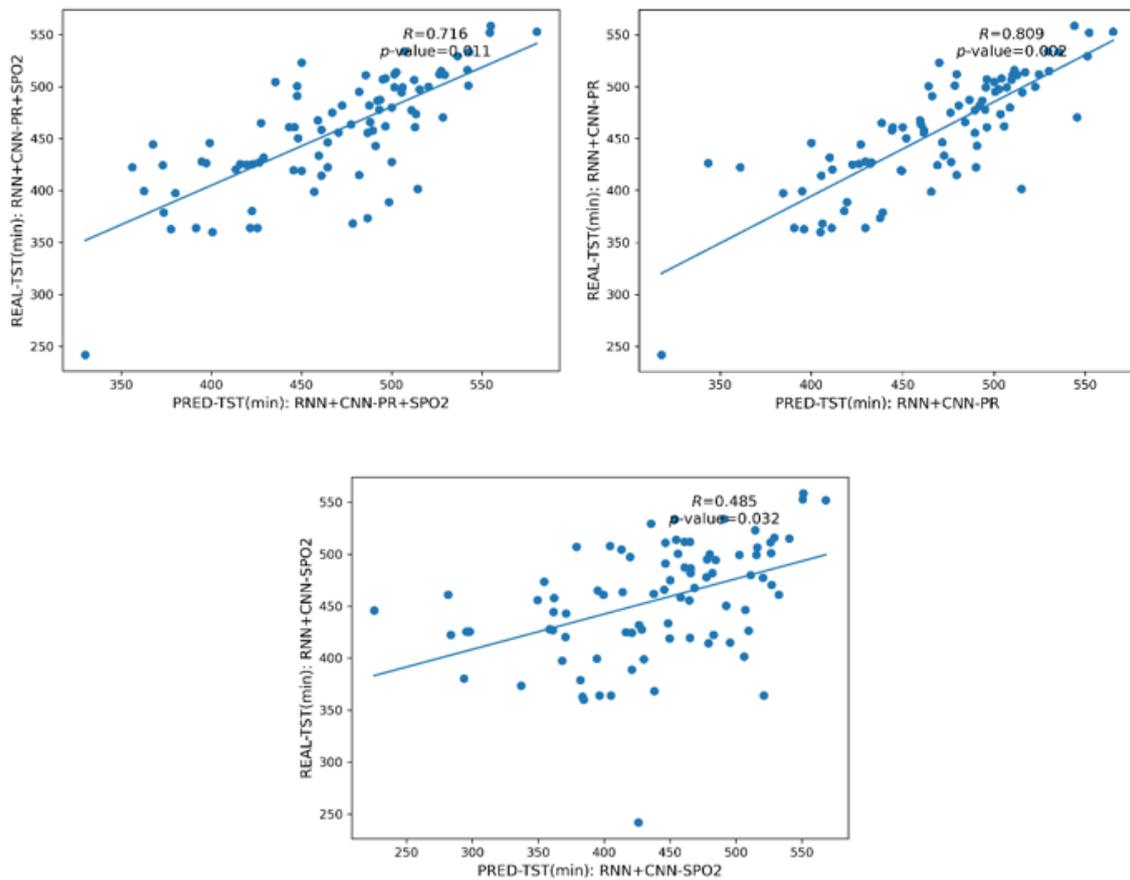


Figura 5.7: Gráfico de dispersión de las señales PR, SpO<sub>2</sub> y PR+ SpO<sub>2</sub> para el TST obtenidos con CNN + RNN.

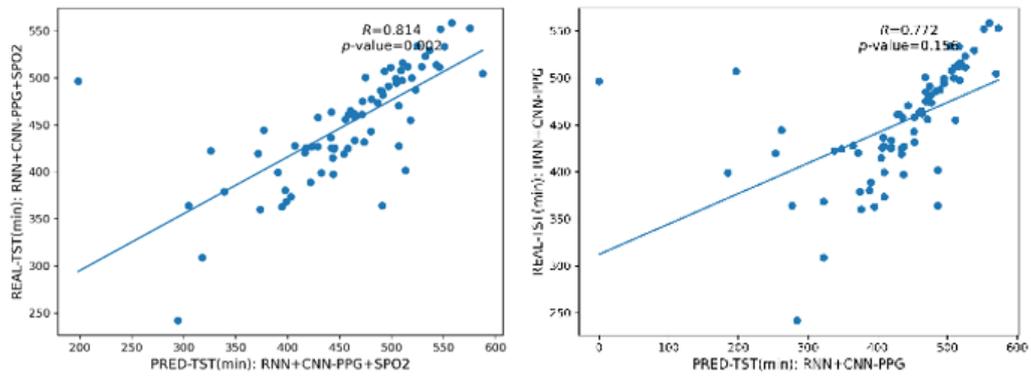


Figura 5.8: Gráfico de dispersión de las señales PPG y PPG+ SpO<sub>2</sub> para el TST obtenidos con CNN + RNN.

## Capítulo 6: DISCUSIÓN

---

En este TFG se ha trabajado con la idea de utilizar CNNs y RNNs como métodos para la detección automática de las fases del sueño a partir de señales de pulsioximetría. A partir de la clasificación de fases del sueño de ambas arquitecturas, se han estimado los parámetros polisomnográficos para cada sujeto. En el apartado anterior se han expuesto los resultados y en este se va a proceder con su análisis e interpretación.

En primer lugar, se discuten los resultados obtenidos en la clasificación de fases del sueño de ambas arquitecturas para las 5 combinaciones de señales diferentes: PR+ SpO<sub>2</sub>, PR, SpO<sub>2</sub> (1Hz), PPG+ SpO<sub>2</sub> Y PPG. Luego, se discuten los resultados obtenidos en la estimación de parámetros polisomnográficos. Posteriormente, se realiza una comparación con los estudios analizados en la revisión del estado del arte y, por último, se comentan las limitaciones de este trabajo realizado.

### 6.1. CLASIFICACIÓN DE FASES DEL SUEÑO

La primera idea a destacar dentro de la detección automática de las fases del sueño (W, N1, N2, N3 y REM) es que los resultados de las arquitecturas CNN y CNN+RNN son bastante aceptables en términos de precisión, Kappa, MF1 y *per-class F1-score*. Hay que distinguir las combinaciones PR+ SpO<sub>2</sub> y PPG+SpO<sub>2</sub> por separado. Sobre la primera podemos destacar que la mejor opción es la combinación de ambas señales, siendo SpO<sub>2</sub> por separado la que peores resultados ofrece, algo lógico por la naturaleza de la señal. En este sentido, la señal de PR es mejor que la de SpO<sub>2</sub> debido a que contiene información sobre los cambios del sistema nervioso autónomo (SNA) relacionado con las fases del sueño, pero la señal de SpO<sub>2</sub> aporta información complementaria sobre las desaturaciones de oxígeno asociadas a los eventos apneicos [49]. Respecto a la opción PPG+SpO<sub>2</sub>, puede parecer sorprendente que se obtengan peores resultados que con la señal de PPG sola, pero esto puede deberse a la suma de dos razones: (i) la señal de SpO<sub>2</sub> se obtiene a partir de la señal PPG a dos longitudes de onda diferentes, por lo que es bastante probable que la información de la señal de SpO<sub>2</sub> sea redundante con la de PPG (ii) problemas de solapamiento por diferencias notables entre los rangos de normalización de ambas señales, lo cual debería de analizarse en detalle en el futuro. A la vista de los resultados de las arquitecturas CNN y CNN+RNN, se han obtenido resultados ligeramente mejores con PPG que con PR+ SpO<sub>2</sub>. Esto puede deberse, como se ha mencionado, a que las señales de PR y SpO<sub>2</sub> se derivan de la PPG. Hay que tener en cuenta que las señales de PR y SpO<sub>2</sub> tienen la ventaja de que pueden ser registradas y almacenadas en pulsioxímetros portátiles, lo que facilita su uso en la práctica clínica en entornos no supervisados. Otro aspecto importante a destacar es que estos dos enfoques han utilizado un número diferente de sujetos de test, por lo que haría falta repetir las pruebas con los mismos sujetos para poder contrastarlo.

Otra idea importante es la comparación del rendimiento entre las dos arquitecturas utilizadas, CNN y CNN+RNN. Los mejores resultados para la primera arquitectura son una precisión de 0,6546 y un Kappa de 0,5394, mientras que para la segunda son una precisión de 0,6879 y un Kappa de 0,587, ambas para la señal de PPG. Podemos apreciar como los resultados con mejores con la arquitectura CNN+RNN.

Por último, podemos apreciar en la optimización de hiperparámetros varias ideas:

- En la CNN, el efecto del solapamiento hace que mejoren los resultados al aumentar las épocas anteriores y posteriores a solapar.
- En la CNN, los otros hiperparámetros tienen una influencia menor en la precisión del conjunto de validación, lo cual destaca la robustez de esta arquitectura.
- En la RNN, los hiperparámetros variados tienen poca influencia en la precisión en validación, indicando también la robustez de esta arquitectura.

## 6.2. ESTIMACIÓN DE PARÁMETROS POLISOMNOGRÁFICOS

En cuanto a los parámetros polisomnográficos se va a destacar el ICC con el que mejor se estiman las fases del sueño.

- N1: Los valores del ICC en la CNN están en el rango 0,0180-0,2547 mientras que para la arquitectura CNN+RNN están en un rango de 0,0159 a 0,3295 siendo la señal de PR la que mejores resultados obtiene.
- N2: Los valores del ICC en la CNN están en el rango 0,1224-0,4053 mientras que para la arquitectura CNN+RNN están en un rango de 0,2652 a 0,3403 siendo la señal de PR+SpO<sub>2</sub> la que mejores resultados obtiene.
- N3: Los valores del ICC en la CNN están en el rango -0,0213-0,1718 mientras que para la arquitectura CNN+RNN están en un rango de 0,0779 a 0,2 siendo la señal de PPG la que mejores resultados obtiene.
- REM: Los valores del ICC en la CNN están en el rango 0,0084-0,5098 mientras que para la arquitectura CNN+RNN están en un rango de -0,0664 a 0,4437 siendo la señal de PPG la que mejores resultados obtiene.
- TST: Los valores del ICC en la CNN están en el rango 0,51-0,7931 mientras que para la arquitectura CNN+RNN están en un rango de 0,3780 a 0,8456 siendo la señal de PR la que mejores resultados obtiene.

En general, en ambas estructuras, los mejores resultados se obtienen con las señales de PR+SpO<sub>2</sub> Y PPG. La utilidad clínica de esta estimación es que las señales que más fácilmente se pueden obtener, mediante dispositivos portátiles, son las que mejores se estiman con estos parámetros, la de PR+SpO<sub>2</sub>.

Por otra parte, teniendo en cuentas las 5 fases del sueño, podemos comprobar que los mejores resultados se obtienen con el TST y los peores resultados con la fase N1.

### 6.3. COMPARACIÓN CON OTROS ESTUDIOS

La tabla 6.1 muestra algunas de las principales características y resultados de los estudios estado del arte de métodos de detección automática de las fases del sueño mediante *deep learning*. Una comparación directa y exhaustiva entre los resultados de los distintos estudios no es posible, ya que no estos estudios usan diferentes bases de datos, rangos de edad, número de pacientes o tipo de señales, entre otras características. Asimismo, no todos los estudios diferencian entre las 5 fases del sueño, por lo que los valores de precisión y kappa no son directamente comparables. Los mejores resultados obtenidos en este estudio para la primera arquitectura (CNN) son una precisión de 0,6546 y un Kappa de 0,5394, mientras que para la segunda (CNN+RNN) son una precisión de 0,6879 y un Kappa de 0,587, ambas para la señal de PPG.

Para realizar una comparación, podemos comparar los estudios que han trabajado con 5 fases del sueño con EEG, como el estudio de Sundar [34], que obtiene resultados de precisión de 90%, o de ECG, como el estudio de Sturburic [51], que obtiene unos resultados de precisión de 40%. Podemos observar en otros estudios, como el de Casal [56] o el de Korkalainen [36] que los resultados mejoran si las fases del sueño a clasificar se reducen. Por último, el estudio a destacar es el de Henri Korkalainen [36], utilizando la señal de PPG y dando unos valores de precisión y kappa que hemos conseguido mejorar con la clasificación de las 5 fases del sueño. En ese estudio se obtiene una precisión de 0,641 y un Kappa de 0,51 mientras que en este TFG se ha conseguido mejorar hasta una precisión de 0,6879 y un Kappa de 0,587. Además, en este TFG no solo se utiliza la señal de PPG, sino que también se han obtenido resultados similares usando las señales de PR y la de SpO<sub>2</sub>, que tienen la ventaja de ser adquiridas y almacenadas por pulsioxímetros portátiles para un posterior procesamiento.

Este estudio permite obtener una gran idea sobre la viabilidad del estudio de señales de pulsioximetría para la detección de las fases del sueño. Esto nos puede aportar a futuro un sistema económico y no invasivo para tratar de diagnosticar trastornos del sueño y posibles patologías, así como su posible tratamiento.

Tabla 6.1: Estudios consultados en la revisión del estado del arte.

ARTÍCULO	SEÑAL	FASES	SUJETOS (N)	EDAD	MÉTODO	RESULTADOS
Minfang et al. (2022) [50].	SINGLE-LEAD ECG	W + R + LS(S1+S2) + DS(S3+S4)	5793 + 2651 + 2056	ADULTOS	DEEP ADAPTATION NETWORK	PRECISIÓN = 79%, 75% Y 80% KAPPA = 0.75, 0.66 Y 0.70
Ebrahim et al. (2021) [33].	EEG	W + S1 + S2 + S3/S4 + R	N/A	ADULTOS	CNN + TCNN + CONDITIONAL RANDOM FIELD LAYER	PRECISIÓN = 82.46% KAPPA = 0.76
Stuburic et al. (2020) [51].	ECG + RESPIRACIÓN + MOVIMIENTO	- W + R + LS(S1+S2) + DS(S3+S4) - W + R + S1 + S2 + (S3 + S4)	23 (18 ENTREN)	38.5 +- 14	CNN + LONG-SHORT TERM MEMORY NETWORK	PRECISIÓN = 55% Y 40%
Jadhav et al. (2020) [52].	EEG	W + R + S1 + S2 + S3/S4	42	ADULTOS	CNN + CWT	PRECISIÓN = 84%
Yang et al. (2021) [53].	EEG	W + S1 + S2 + SWS + R	20 + 20	25-35 20-65	1D-CNN-HIDDEN MAKOV MODEL	PRECISIÓN = 83.98% Y 81.68% KAPPA = 0.78 Y 0.74
Werth et al. (2019) [35].	ECG + HRV	ACTIVE + QUIET + INTERMEDIATE + WAKE + CARETAKING	34	NEONATOS PREMATUROS	CNN + RNN	KAPPA = 0.43 +- 0.008
Sundar et al. (2021) [34].	EEG	W + S1 + S2 + S3 + R	197	ADULTOS	CNN (DECODIF) + RNN (CODIF)	PRECISIÓN = 90.7 %
Krauss et al. (2021) [54].	EEG	W + S1 + S2 + S3 + R	68	32.5 + 11.5	DPN	N/A
Yan et al. (2021) [55].	EEG + EOG + EMG + ECG	W + S1 + S2 + S3 + R	100 + 99 + 19	46.86 +- 4.22 51 +- 16 28.74 +- 2.99	CNN	PRECISIÓN: 87%, 86 % Y 86% KAPPA= 0.81, 0.82 Y 0.81
Casal et al. (2021) [56].	HRV + SpO <sub>2</sub>	W + (S1/S2/S3/R)	5000 (2500 ENTRENAR RED)	ADULTOS	RNN	PRECISIÓN = 90.13% KAPPA= 0.74
Korkalainen et al. (2020) [36].	PPG	- W + NR + R - W + S1/S2 + S3 + R - W + S1 + S2 + S3 + R	894	44.7 – 65.8	CNN + RNN	PRECISIÓN: 80.1%, 68.5% Y 64.1 % KAPPA= 0.65, 0.54 Y 0.51
Vaquerizo et al. (2022) [37].	PPG	W + REM + NREM	366	5-8	CNN	PRECISIÓN= 78.2% KAPPA=0.57

## 6.4. LIMITACIONES

Para finalizar el capítulo de discusión se van a detallar las limitaciones presentes en el trabajo realizado.

La primera limitación de este estudio es que se ha trabajado con un número distinto de sujetos con la combinación PPG+SpO<sub>2</sub> (y PPG) que con PR+SpO<sub>2</sub> (y PR y SpO<sub>2</sub> por separado), con el objetivo de trabajar con el mayor número de señales posibles

Otra limitación es el rango de edad de los sujetos, ya que solo contamos con niños entre 5 y 10 años, lo que puede implicar un ligero cambio en los resultados para niños de edades fuera del rango. Además, la base de datos solo tiene sujetos con un AHI>1, es decir, no hay sujetos sin AOS.

La siguiente limitación es que no se ha proporcionado una explicación e interpretación de las características de pulsioximetría que las arquitecturas CNN y CNN+RNN han aprendido en cada relación a cada fase del sueño.

Otra limitación es el *inter-scorer variability* [57] puede afectar a las fases clasificadas. Este problema se debe a las épocas que son difíciles de clasificar.



# Capítulo 7: CONCLUSIONES Y LÍNEAS FUTURAS

---

En este trabajo se ha profundizado en el uso conjunto de CNNs con RNNs para la detección automática del sueño en señales menos comunes para ello, las señales de pulsioximetría. Además, las bases de datos empleadas son de niños, algo novedoso respecto a los trabajos previos publicados.

Primero, se van a indicar las aportaciones al campo de investigación de este TFG. Posteriormente se van a enunciar las conclusiones conseguidas y se va a terminar con las líneas futuras de investigación sobre el tema tratado.

## 7.1. CONTRIBUCIONES

Las principales contribuciones que este TFG aporta al campo de la investigación sobre los métodos y señales válidas para la detección automática de las fases del sueño se pueden resumir en tres puntos principales:

- En primer lugar, el uso de una base de datos en la que los sujetos son niños con el objetivo de poder obtener resultados para este tipo de personas, ya que los resultados existentes en su mayoría son para adultos. Esto es importante porque en la mayoría de casos los niños difieren de los adultos, tanto en características de las señales como en comportamientos y posibles patologías.
- El uso de señales de pulsioximetría, con unos resultados más que aceptables, en lugar de las señales típicas de EEG o ECG, nos permite obtener un método más sencillo, de menor coste y menos invasivo para la detección de las fases del sueño y, por lo tanto, de las posibles patologías asociadas a él. Además, realizamos una comparación del rendimiento de distintas configuraciones de señales de entrada.
- La estimación de parámetros polisomnográficos en niños, con el correspondiente análisis de la utilidad práctica de estas estimaciones.
- Por último, el empleo de *deep learning* como método de clasificación juntando dos tipos de redes neuronales, CNN Y RNN, nos permite comprobar que empleándolas simultáneamente se pueden obtener mejores resultados. Además, se ha introducido cierto *dropout* en capas convolucionales, algo no muy usado en los estudios publicados por otros autores.

## 7.2. CONCLUSIONES

Las conclusiones finales obtenidas se exponen en los siguientes puntos:

- Los métodos de detección y clasificación automática de las fases del sueño se presentan como una gran opción que aporta objetividad y ahorra tiempo y dinero.
- El uso de métodos de *deep learning* permite alcanzar un resultado notable en la detección de fases del sueño a partir de las señales de pulsioximetría. El mejor resultado conseguido fue una precisión de 68,0% y un Kappa de 0,5761 con la señal de PPG por separado.

- Estos métodos se presentan como una gran opción para detectar fases del sueño en estudios de pulsioximetría para el diagnóstico de la AOS infantil. Se ha conseguido demostrar que las señales de pulsioximetría estudiadas (PR, SpO<sub>2</sub> y PPG) contienen información útil para detectar las fases del sueño y pueden ser procesadas mediante redes neuronales para poder obtener una clasificación aceptable de las fases del sueño.
- Las señales que mejor se comportan son las de PPG y PPG+ SpO<sub>2</sub>. También se obtienen buenos resultados con la combinación PR+SpO<sub>2</sub>, algo importante debido a que son las señales que podemos obtener con dispositivos portátiles comunes de medida.
- Para realizar el preprocesado de las señales de las bases de datos es absolutamente imprescindible homogeneizar las señales a una frecuencia de muestreo común y realizar una normalización, ya que cualquier variación puede afectar y superponer información o deformarla, como hemos podido comprobar especialmente en la señal de PPG+SpO<sub>2</sub>.
- Se ha comprobado la importancia de dividir los sujetos en los diferentes grupos. Es clave para las redes neuronales y el *deep learning* tener un buen grupo de entrenamiento para poder mejorar los resultados a obtener mediante los de validación y test.
- Se ha comprobado que el uso únicamente de CNN consigue buenos resultados, pero que al incluir una RNN se consiguen mejores resultados.
- Un gran número de capas convolucionales puede producir overfitting y empeorar los resultados. Asimismo, se ha comprobado en la optimización de parámetros que no siempre más es mejor, esto se puede observar con el *dropout* o con el número de unidades GRU.
- Hay fases del sueño que se detectan mejor que otras, siendo la N1 la más complicada y W la más clara, debido a sus características. Los resultados obtenidos se han conseguido diferenciando 5 fases, si hubiéramos juntado fases de sueño ligero y sueño profundo se hubieran mejorado los resultados.

Como observación final, se puede concluir que este TFG ha servido para comprobar que el uso de CNNs y RNNs en señales de pulsioximetría permite detectar correctamente las fases del sueño y es un método que merece la pena estudiar, investigar y mejorar.

### 7.3. LÍNEAS FUTURAS

En futuras investigaciones o trabajos se podría indagar o estudiar en más profundidad ciertos aspectos que han estado limitados por el tiempo en este TFG.

Respecto a la base de datos se podrían obtener sujetos no-AOS (AHÍ<1), de un rango un poco más amplio de edad para conseguir más ejemplos de entrenamiento para las redes neuronales y de diferentes rangos de severidad de la AOS. Además, se podrían obtener señales un poco más limpias de ruido para poder procesar la información mejor.

Respecto al preprocesado se podría conseguir que la señal de PPG+SpO<sub>2</sub> estén normalizadas a unos valores similares ya que hemos visto que en ciertos resultados la información de ambas se ha podido superponer haciendo que la detección empeore ligeramente, como podemos comprobar con los resultados de precisión y Kappa, respecto a la señal de PPG aislada.

Se podría continuar el trabajo de investigación con otras técnicas a mayores de la CNN y la RNN como pueden ser los Transformers [58], un tipo de red neuronal que es capaz de aprender el contexto y significado mediante el seguimiento de relaciones en datos secuenciales. Es un método pensado para continuar este trabajo con el TFM. Otra opción es el uso de técnicas de *explicable artificial intelligence* (XAI) para descubrir qué características de las señales de pulsioximetría están relacionadas con cada fase.

Finalmente, habría que calcular el coste de aplicar estas técnicas a un modelo portátil de captación de señales de pulsioximetría para comprobar su viabilidad y validar la propuesta en señales captadas en un entorno domiciliario con estos dispositivos, para comprobar el rendimiento.



## REFERENCIAS

- [1] Bronzino, J. D. (2013). *The Biomedical Engineering Handbook. Medical Devices and Systems* (Vol. 53).
- [2] Sörnmo, L. & Laguna, P. (2005). In *Bioelectrical Signal Processing in Cardiac and Neurological Applications* (p. iv). Elsevier. <https://doi.org/10.1016/B978-0-12-437552-9.50013-1>
- [3] Poza Crespo, J. (2008). *Análisis tiempo-frecuencia de la actividad magnetoencefalográfica espontánea en la enfermedad de Alzheimer*. <https://doi.org/10.35376/10324/136>
- [4] Marcus, C. L., Brooks, L. J., Ward, S. D., Draper, K. A., Gozal, D., Halbower, A. C., Jones, J., Lehmann, C., Schechter, M. S., Sheldon, S., Shiffman, R. N., & Spruyt, K. (2012). Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome. *Pediatrics*, *130*(3), e714–e755. <https://doi.org/10.1542/peds.2012-1672>
- [5] Piera Fernández, M. (2001). Síndrome de la apnea obstructiva del sueño. Descripción y tratamiento. *Farmacia Profesional*, *15*(3), 62–69.
- [6] Katz, E. S., Mitchell, R. B., & D'Ambrosio, C. M. (2012). Obstructive Sleep Apnea in Infants. *American Journal of Respiratory and Critical Care Medicine*, *185*(8), 805–816. <https://doi.org/10.1164/rccm.201108-1455CI>
- [7] Tan, H.-L., Gozal, D., Ramirez, H. M., Bandla, H. P. R., & Kheirandish-Gozal, L. (2014). Overnight Polysomnography versus Respiratory Polygraphy in the Diagnosis of Pediatric Obstructive Sleep Apnea. *Sleep*, *37*(2), 255–260. <https://doi.org/10.5665/sleep.3392>
- [8] del Campo, F., Crespo, A., Cerezo-Hernández, A., Gutiérrez-Tobal, G. C., Hornero, R., & Álvarez, D. (2018). Oximetry use in obstructive sleep apnea. *Expert Review of Respiratory Medicine*, *12*(8), 665–681. <https://doi.org/10.1080/17476348.2018.1495563>
- [9] Phan, H., Andreotti, F., Cooray, N., Chen, O. Y., & de Vos, M. (2019). SeqSleepNet: End-to-End Hierarchical Recurrent Neural Network for Sequence-to-Sequence Automatic Sleep Staging. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *27*(3), 400–410. <https://doi.org/10.1109/TNSRE.2019.2896659>
- [10] Boostani, R., Karimzadeh, F., & Nami, M. (2017). A comparative review on sleep stage classification methods in patients and healthy individuals. *Computer Methods and Programs in Biomedicine*, *140*, 77–91. <https://doi.org/10.1016/j.cmpb.2016.12.004>
- [11] Iber, C., Ancoli-Israel, S., Chesson, A. and Quan, S.F. (2007) The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications. American Academy of Sleep Medicine, Westchester.
- [12] Cerutti, S & Marchesi, C (2011). *Advanced Methods of Biomedical Signal Processing*.

- [13] Penzel, T., Zhang, X., & Fietze, I. (2013). Inter-scoring Reliability between Sleep Centers Can Teach Us What to Improve in the Scoring Rules. *Journal of Clinical Sleep Medicine*, 09(01), 89–91. <https://doi.org/10.5664/jcsm.2352>
- [14] Chan, E. D., Chan, M. M., & Chan, M. M. (2013). Pulse oximetry: Understanding its basic principles facilitates appreciation of its limitations. *Respiratory Medicine*, 107(6), 789–799. <https://doi.org/10.1016/j.rmed.2013.02.004>
- [15] A review on wearable photoplethysmography sensors and their potential future applications in health care  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6426305/#:~:text=A%20typical%20PPG%20device%20contains,proportional%20to%20blood%20volume%20variations.>
- [16] Mohan, P. M., Nisha, A. A., Nagarajan, V., & Jothi, E. S. J. (2016). Measurement of arterial oxygen saturation (SpO<sub>2</sub>) using PPG optical sensor. *2016 International Conference on Communication and Signal Processing (ICCSP)*, 1136–1140. <https://doi.org/10.1109/ICCSP.2016.7754330>
- [17] Aggarwal, C. C. (2018). Convolutional Neural Networks. In *Neural Networks and Deep Learning* (pp. 315–371). Springer International Publishing. [https://doi.org/10.1007/978-3-319-94463-0\\_8](https://doi.org/10.1007/978-3-319-94463-0_8)
- [18] ¿Cuál es la diferencia entre el machine learning y el deep learning?  
<https://blog.bismart.com/diferencia-machine-learning-deep-learning>
- [19] Redline, S., Amin, R., Beebe, D., Chervin, R. D., Garetz, S. L., Giordani, B., Marcus, C. L., Moore, R. H., Rosen, C. L., Arens, R., Gozal, D., Katz, E. S., Mitchell, R. B., Muzumdar, H., Taylor, H. G., Thomas, N., & Ellenberg, S. (2011). The Childhood Adenotonsillectomy Trial (CHAT): Rationale, Design, and Challenges of a Randomized Controlled Trial Evaluating a Standard Surgical Procedure in a Pediatric Population. *Sleep*, 34(11), 1509–1517. <https://doi.org/10.5665/sleep.1388>
- [20] Machine Learning | Coursera. <https://www.coursera.org/learn/machine-learning>.
- [21] Marcus, C. L., Brooks, L. J., Ward, S. D., Draper, K. A., Gozal, D., Halbower, A. C., Jones, J., Lehmann, C., Schechter, M. S., Sheldon, S., Shiffman, R. N., & Spruyt, K. (2012). Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome. *Pediatrics*, 130(3), e714–e755. <https://doi.org/10.1542/peds.2012-1672>
- [22] Síndrome de apnea obstructiva del sueño.  
<http://publicacionesmedicina.uc.cl/Boletin/20072/SindromeApnea.pdf>
- [23] Loh, H. W., Ooi, C. P., Vicnesh, J., Oh, S. L., Faust, O., Gertych, A., & Acharya, U. R. (2020). Automated Detection of Sleep Stages Using Deep Learning Techniques: A Systematic Review of the Last Decade (2010–2020). *Applied Sciences*, 10(24), 8963. <https://doi.org/10.3390/app10248963>
- [24] Rosenblatt, F. (1960). Perceptron Simulation Experiments. *Proceedings of the IRE*, 48(3), 301–309. <https://doi.org/10.1109/JRPROC.1960.287598>

- [25] Principe, J. C., Gala, S. K., & Chang, T. G. (1989). Sleep staging automaton based on the theory of evidence. *IEEE Transactions on Biomedical Engineering*, 36(5), 503–509. <https://doi.org/10.1109/10.24251>
- [26] Ian Witten, E. F. M. H. C. P. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*.
- [27] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- [28] Xanthopoulos, P., Pardalos, P. M., & Trafalis, T. B. (2013). *Robust Data Mining*. Springer New York. <https://doi.org/10.1007/978-1-4419-9878-1>
- [29] Peker, M. (2016). An efficient sleep scoring system based on EEG signal using complex-valued machine learning algorithms. *Neurocomputing*, 207, 165–177. <https://doi.org/10.1016/j.neucom.2016.04.049>
- [30] Sors, A., Bonnet, S., Mirek, S., Vercueil, L., & Payen, J.-F. (2018). A convolutional neural network for sleep stage scoring from raw single-channel EEG. *Biomedical Signal Processing and Control*, 42, 107–114. <https://doi.org/10.1016/j.bspc.2017.12.001>
- [31] Redes Neuronales Recurrentes. <https://torres.ai/redes-neuronales-recurrentes/#:~:text=Las%20redes%20neuronales%20recurrentes%2C%20o,neuronales%20vistas%20en%20cap%C3%ADtulos%20anteriores>.
- [32] Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37-46.
- [33] Khalili, E., & Mohammadzadeh Asl, B. (2021). Automatic Sleep Stage Classification Using Temporal Convolutional Neural Network and New Data Augmentation Technique from Raw Single-Channel EEG. *Computer Methods and Programs in Biomedicine*, 204. <https://doi.org/10.1016/j.cmpb.2021.106063>
- [34] Sundar, G. N., Narmadha, D., Anton Jone, A. A., Sagayam, K. M., Dang, H., & Pomplun, M. (2021). Automated sleep stage classification in sleep apnoea using convolutional neural networks. *Informatics in Medicine Unlocked*, 26. <https://doi.org/10.1016/j.imu.2021.100724>
- [35] Werth, J., Radha, M., Andriessen, P., Aarts, R. M., & Long, X. (2020). Deep learning approach for ECG-based automatic sleep state classification in preterm infants. *Biomedical Signal Processing and Control*, 56. <https://doi.org/10.1016/j.bspc.2019.101663>
- [36] Korkalainen, H., Aakko, J., Duce, B., Kainulainen, S., Leino, A., Nikkonen, S., Afara, I. O., Myllymaa, S., Töyräs, J., & Leppänen, T. (2020). Deep learning enables sleep staging from photoplethysmogram for patients with suspected sleep apnea. *Sleep*, 43(11). <https://doi.org/10.1093/sleep/zsaa098>

- [37] Vaquerizo-Villar F, Alvarez D, Kraemer JF, Wessel N, Gutierrez-Tobal GC, Calvo E, Del Campo F, Kheirandish-Gozal L, Gozal D, Penzel T, Hornero R. Automatic Sleep Staging in Children with Sleep Apnea using Photoplethysmography and Convolutional Neural Networks. *Annu Int Conf IEEE Eng Med Biol Soc.* 2021 Nov;2021:216-219. doi: 10.1109/EMBC46164.2021.9629995. PMID: 34891275.
- [38] Huttunen, R., Leppänen, T., Duce, B., Oksenberg, A., Myllymaa, S., Töyräs, J., & Korkalainen, H. (2021). Assessment of obstructive sleep apnea-related sleep fragmentation utilizing deep learning-based sleep staging from photoplethysmography. *Sleep*, 44(10). <https://doi.org/10.1093/sleep/zsab142>
- [39] Nouri, S. (2012). Supercharge Your Data Science Career!
- [40] Raschka, S., Mirjalili, V. (2017). Python Machine Learning. Packt Publishing Ltd.
- [41] Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. In *Neural Networks* (Vol. 61, pp. 85–117). Elsevier Ltd. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [42] Convolutional Neural Network - Long - Medium. <https://medium.com/@Aj.Cheng/convolutional-neural-network-d9f69e473feb>
- [43] Srivastava, N., Hinton, G., Krizhevsky, A., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. In *Journal of Machine Learning Research* (Vol. 15).
- [44] Accuracy, Precision, Recall & F1-Score – Python Examples. <https://vitalflux.com/accuracy-precision-recall-f1-score-python-example/>
- [45] Abaira, V. (2001). El índice kappa. *SEMERGEN - Medicina de Familia*, 27(5), 247–249. [https://doi.org/10.1016/s1138-3593\(01\)73955-x](https://doi.org/10.1016/s1138-3593(01)73955-x)
- [46] The F1 score. <https://towardsdatascience.com/the-f1-score-bec2bbc38aa6>
- [47] Coeficiente de correlación intraclase. [https://metodo.uab.cat/arxiu/material/general/El\\_coeficiente\\_de\\_correlacion\\_intraclase.pdf](https://metodo.uab.cat/arxiu/material/general/El_coeficiente_de_correlacion_intraclase.pdf)
- [48] ¿Qué es el coeficiente de correlación de Spearman? <https://www.questionpro.com/blog/es/coeficiente-de-correlacion-de-spearman/>
- [49] Vaquerizo-Villar, F., Penzel, T., Álvarez, D., Kheirandish-Gozal, L., Hornero, R., Gutiérrez-Tobal, G. C., & Gozal, D. (n.d.). *A convolutional neural network to classify sleep stages in pediatric sleep apnea from pulse oximetry signals.*
- [50] Tang, M., Zhang, Z., He, Z., Li, W., Mou, X., Du, L., Wang, P., Zhao, Z., Chen, X., Li, X., Chang, H., & Fang, Z. (2022). Deep adaptation network for subject-specific sleep stage classification based on a single-lead ECG. *Biomedical Signal Processing and Control*, 75. <https://doi.org/10.1016/j.bspc.2022.103548>

- [51] Stuburic, K., Gaiduk, M., & Seepold, R. (2020). A deep learning approach to detect sleep stages. *Procedia Computer Science*, 176, 2764–2772. <https://doi.org/10.1016/j.procs.2020.09.280>
- [52] Jadhav, P., Rajguru, G., Datta, D., & Mukhopadhyay, S. (2020). Automatic sleep stage classification using time–frequency images of CWT and transfer learning using convolution neural network. *Biocybernetics and Biomedical Engineering*, 40(1), 494–504. <https://doi.org/10.1016/J.BBE.2020.01.010>
- [53] Yang, B., Zhu, X., Liu, Y., & Liu, H. (2021). A single-channel EEG based automatic sleep stage classification method leveraging deep one-dimensional convolutional neural network and hidden Markov model. *Biomedical Signal Processing and Control*, 68. <https://doi.org/10.1016/j.bspc.2021.102581>
- [54] Krauss, P., Metzner, C., Joshi, N., Schulze, H., Traxdorf, M., Maier, A., & Schilling, A. (2021). Analysis and visualization of sleep stages based on deep neural networks. *Neurobiology of Sleep and Circadian Rhythms*, 10. <https://doi.org/10.1016/j.nbscr.2021.100064>
- [55] Yan, R., Li, F., Zhou, D. D., Ristaniemi, T., & Cong, F. (2021). Automatic sleep scoring: A deep learning architecture for multi-modality time series. *Journal of Neuroscience Methods*, 348. <https://doi.org/10.1016/j.jneumeth.2020.108971>
- [56] Casal, R., di Persia, L. E., & Schlotthauer, G. (2021). Classifying sleep–wake stages through recurrent neural networks using pulse oximetry signals. *Biomedical Signal Processing and Control*, 63. <https://doi.org/10.1016/j.bspc.2020.102195>
- [57] Younes, M., Raneri, J., & Hanly, P. (2016). Staging sleep in polysomnograms: Analysis of inter-scorer variability. *Journal of Clinical Sleep Medicine*, 12(6), 885–894. <https://doi.org/10.5664/jcsm.5894>
- [58] ¿Qué es un Modelo Transformer? <https://la.blogs.nvidia.com/2022/04/19/que-es-un-modelo-transformer/>