



UTN.BA

UNIVERSIDAD TECNOLÓGICA NACIONAL
FACULTAD REGIONAL BUENOS AIRES

Grupo 12

Augusto Javier Santini
asantini@est.frba.utn.edu.ar
Legajo: 149.197-0

Proyecto Final

**Detector de complejos QRS en base a
redes neuronales**

2019

Tabla de contenidos

1. Introducción	3
2. Marco Teórico	3
2.1. Señales electrocardiográficas y el procesamiento de las mismas	3
2.1.1. Anatomía y actividad eléctrica del corazón	4
2.1.2. Electrocardiograma (ECG)	6
2.1.3. Características del ECG	7
2.1.4. Ruido en el ECG	8
2.1.5. Procesamiento de ECG	9
2.1.6. Detección del complejo QRS	9
2.2. Redes neuronales	10
2.2.1. Generalidades	10
2.2.2. Conexiones y analogías biológicas	10
2.2.3. Arquitectura de una red neuronal	11
2.2.4. Tipos de capas neuronales básicas	11
2.2.5. Tipos de funciones de activación básicas	12
2.2.6. Proceso de aprendizaje de la red neuronal	13
3. Materiales	14
3.1. Registros de ECG de seres humanos	14
3.2. Registros de ECG de roedores	15
3.3. Lenguaje de programación y APIs utilizadas	16
4. Metodología	17
4.1. Preparación de los datos a utilizar	17
4.1.1. Canales a utilizar de cada registro	17
4.1.2. Resampleo, filtrado y ventana temporal utilizada	17
4.1.3. Aumentación de datos	17
4.1.4. Sobreentrenamiento	18
4.1.5. Separación de datos para cada set	19
4.2. Arquitecturas neuronales planteadas	20
4.2.1. Primer arquitectura	20
4.2.2. Modificaciones realizadas a la arquitectura planteada	21

4.2.3. Arquitectura final	21
4.3. Método de evaluación del detector	22
4.3.1. Matriz de confusión	22
5. Resultados	23
5.1. Resultados en ECG de seres humanos con ambas arquitecturas sobre los sets de validación	24
5.2. Resultados en ECG de seres humanos con la arquitectura final en el set de evaluación	25
5.2.1. Resultados en ECG de roedores	25
5.3. Resultados registros sintéticos	26
5.4. Comparación con otro algoritmo de detección utilizado actualmente	26
5.5. Proceso de entrenamiento para la red neuronal	26
6. Conclusiones	26

1. Introducción

Las enfermedades cardiovasculares son la primer causa de muerte a escala global. Según la Organización Mundial de la Salud, 17.9 millones de personas han muerto en años recientes a causa de enfermedades cardiovasculares, representando un 31 % de la totalidad de las muertes en el mundo [1]. Siendo el tópico de las enfermedades cardiovasculares una problemática actual y de alta gravedad, se trabajó en una problemática particular a partir del procesamiento de señales, en conjunto con otra universidad.

El corazón consta de dos actividades principales, una *mecánica* y una *eléctrica*, siendo la última la representada mediante la señal *ECG*, la trabajada en éste proyecto. En cualquier diagnóstico automático o semi-automático, el primer paso es detectar puntos característicos de la señal ECG siendo uno de los más importantes el *complejo QRS*. Si éstos puntos no fueron bien detectados, todos los pasos subsiguientes arrastrarán el error generado por dicho detector por lo que es una etapa donde la fiabilidad es altamente crítica. A su vez, no sólo se utilizan las detecciones de complejos QRS para diagnósticos automáticos, sino también para propósitos de investigación en el ámbito de la medicina.

A lo largo de éste proyecto se desarrollará un detector de complejos QRS basado en *redes neuronales*, el cual será capaz de producir detecciones fiables tanto en ECG de seres humanos como en ECG de roedores. Se buscó este objetivo ya que en particular se trabajó en conjunto con un grupo de investigación de medicina de la *UNCUYO* (Universidad Nacional de Cuyo) donde trabajan con señales obtenidas de corazones de roedores explantados, pudiendo hacer experimentos y evaluaciones que no podrían hacer en otras circunstancias, ya que no podrían acceder a características tan particulares como la solución en la que se encuentra inmersa el corazón, o mediciones a nivel celular como pueden ser mediciones de *potenciales de acción*.

Se optó por un enfoque basado en redes neuronales, dado a la gran potencia de solucionar problemas complejos con poco conocimiento previo del problema, ya que si bien, las características de ECG de seres humanos están vastamente estudiadas y existen detectores confiables en dicho ámbito, los mismos detectores no funcionan de manera correcta para ECG de roedores, dadas las diferencias en las características de ambas señales.

En el GIBIO (Grupo de Investigación de Bioingeniería), perteneciente a la UTN-FRBA (Universidad Tecnológica Nacional, Facultad Regional Buenos Aires) existen diversos proyectos en conjunto con otros grupos de investigación. En este trabajo se desarrolló un detector de complejos QRS orientado a colaborar con un grupo de investigación perteneciente a la UNCUYO (Universidad Nacional de Cuyo, Mendoza) el cual se dedica a investigación médica en base a experimentos realizados en diversos campos. En particular, en este trabajo se desarrolló un detector de complejos QRS que sea capaz de detectar tanto en ECG de seres humanos como en ECG de roedores. Para tal objetivo se propuso realizar el detector en base a *Redes Neuronales*, dado que las mismas, dada la capacidad suficiente, tienen la capacidad de resolver problemas muy diversos. En particular, se propuso experimentar con bases de datos de ECG pertenecientes a seres humanos, para indagar en la posibilidad del buen funcionamiento del enfoque, y una vez alcanzado un rendimiento aceptable, pasar a los ECG pertenecientes a los roedores.

2. Marco Teórico

2.1. Señales electrocardiográficas y el procesamiento de las mismas

El marco teórico relacionado con la anatomofisiología del corazón, características de las señales electrocardiográficas y teoría de detección de complejos QRS, incluyendo varias de las figuras expuestas a continuación, es sostenido desde el texto bibliográfico *Bioelectrical Signal Processing in Cardiac and Neurological Applications* [2].

2.1.1. Anatomía y actividad eléctrica del corazón

El corazón es un órgano muscular cuya función principal es bombear sangre oxigenada a lo largo del cuerpo. Su anatomía está dividida en dos lados “espejados” los cuales manejan distintos sistemas circulatorios, pero cuyos bombeos están sincronizados de una forma rítmica.

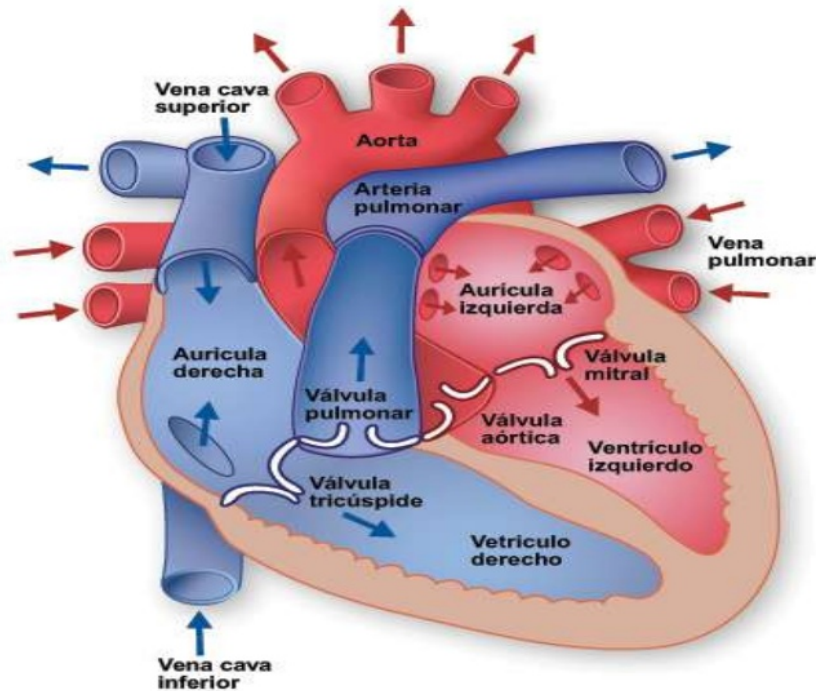


Figura 1: Anatomía del corazón

En la Figura 1 se puede observar la anatomía del corazón. Cada uno de sus lados, consiste de dos cavidades, la *aurícula* por donde entra la sangre, y el *ventrículo* por donde la sangre sale para su circulación. Los dos lados se encuentran separados por una pared muscular llamada *septum*. La dirección del flujo de sangre es controlado por dos válvulas localizadas entre las aurículas y los ventrículos, denominadas válvulas *auriculoventriculares* y otras dos válvulas localizadas entre los ventrículos y las arterias, denominadas válvulas *pulmonares* (lado derecho del corazón) y *aórtica* (lado izquierdo del corazón). La pared externa del corazón es denominada *miocardio* y está principalmente compuesta por células musculares que producen una fuerza mecánica durante la contracción del corazón. Asimismo, el miocardio contiene células musculares especializadas las cuales están conectadas a una red denominada *sistema conductivo* que permite a un impulso eléctrico diseminarse rápidamente a través del corazón. Un *ciclo cardíaco* es iniciado cuando dicho impulso se propaga a través del sistema conductivo. El impulso eléctrico es el evento que dispara la fuerza mecánica, y por consiguiente el evento eléctrico precede al mecánico.

La secuencia de eventos mecánicos que definen un ciclo cardíaco se puede asumir que comienza en la aurícula derecha, donde la sangre es recolectada de todas las venas del cuerpo, exceptuando a las que se encuentran en los pulmones. Cuando la aurícula derecha se contrae, la misma fuerza sangre en el ventrículo derecho. Cuando el ventrículo derecho fue llenado, se contrae y fuerza sangre en los pulmones, momento en el cual el exceso de dióxido de carbono es reemplazado por oxígeno. Las venas pulmonares retornan la sangre ya oxigenada en la aurícula izquierda, la cual luego es vaciada en el ventrículo izquierdo. El ventrículo izquierdo tiene una capacidad de bomba de alta presión, forzando entonces el flujo de sangre hacia todo el cuerpo (exceptuando los pulmones) a través de las arterias.

Cada ciclo cardíaco está compuesto por dos fases, activación y recuperación, a las cuales se refiere en términos eléctricos como *despolarización* y *repolarización*, y en términos mecánicos como *contracción o sístole* y *relajación o diástole*.

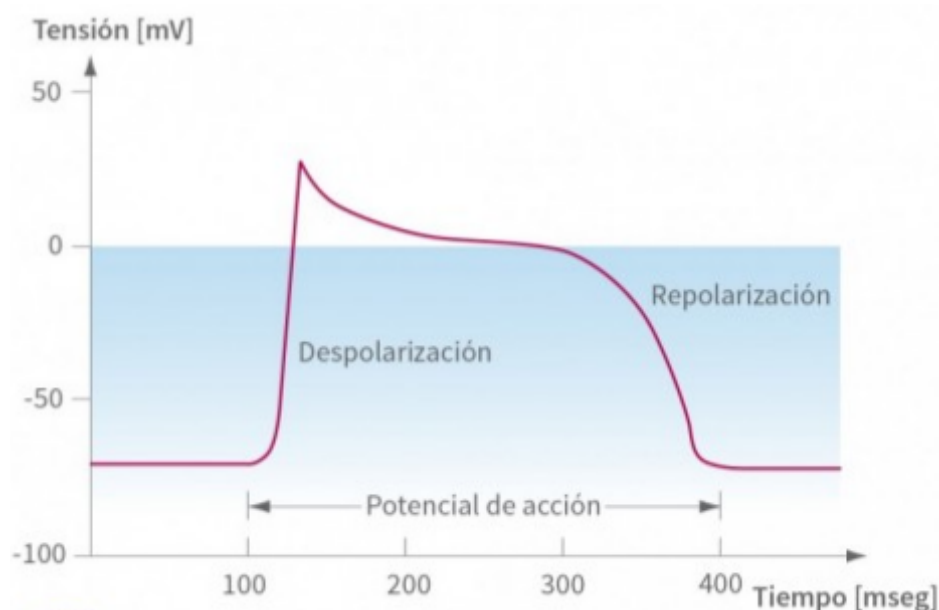


Figura 2: Potencial de acción cardíaco

La despolarización es manifestada por un cambio súbito en el potencial de membrana de la célula y constituye la fase inicial del potencial de acción cardíaco. El cambio súbito de potencial causa una despolarización en las células aledañas, resultando en una dispersión de un impulso eléctrico de célula a célula, a través del miocardio. Inmediatamente después de la despolarización, sigue una repolarización, durante la cual el potencial de membrana recupera gradualmente su estado de reposo, como se puede observar en la Figura 2.

El inicio de un ciclo cardíaco ocurre en un conjunto de células marcapaso que tienen la habilidad de disparar espontáneamente un impulso eléctrico. Se refiere a este conjunto de células como *nódulo sinoauricular* y están situadas en la parte superior de la aurícula derecha. El impulso eléctrico, también denominado *frente de onda*, es entonces propagado a través del sistema de conducción de forma tal de que la contracción tanto auricular como ventricular pueda suceder en el tiempo adecuado. Luego de la activación eléctrica de ambas aurículas, el impulso eléctrico es retrasado en el nódulo auriculoventricular antes de llegar a los ventrículos. Este retardo temporal le da tiempo a la contracción auricular para incrementar el volumen de sangre en los ventrículos antes de que ocurra la contracción ventricular. El retraso en el nódulo auriculoventricular es dado por la baja conducción eléctrica por el tejido muscular en dicha área. El impulso eléctrico entra en la pared que divide ambos ventrículos en el haz de His, único lugar donde están conectadas eléctricamente la aurícula y el nódulo auriculoventricular. El camino eléctrico es entonces dividido en haces de rápida conducción con ramificaciones hacia ambos ventrículos y luego hacia una red de fibras de conducción especializadas, llamadas *fibras de Purkinje*. Como el tamaño de los ventrículos es significativamente grande, es necesario que el impulso eléctrico se propague rápidamente para iniciar una contracción unificada, por lo que la velocidad de conducción en las *fibras de Purkinje* es aproximadamente 80 veces más rápida que la velocidad de conducción del nódulo auriculoventricular.

El nódulo sinoauricular es el marcapasos natural que tiene el corazón, el cual determina el ritmo cardíaco. Como las células del nódulo sinoauricular tienen el ritmo de marcapasos más rápido, todas las demás células las siguen en sincronía. Sin embargo, el ritmo del nódulo sinoauricular no es determinado únicamente por su ritmo de descarga inherente, sino también por información externa, la cual es mediada a través del sistema nervioso autónomo. El balance entre las partes parasimpáticas y simpáticas del sistema nervioso autónomo determinan el ritmo cardíaco, de forma tal que un incremento en la actividad parasimpática reduce el ritmo cardíaco, y un incremento en la actividad simpática incrementa el ritmo cardíaco. Existe una cota superior en el ritmo de descarga, marcada por el tiempo durante el cual una célula está eléctricamente inactiva, denominado *período refractario*, el cual puede variar entre 200ms y 250ms, lo cual se corresponde con un ritmo cardíaco máximo teórico entre 240bpm y 300bpm.

La actividad eléctrica del corazón puede ser caracterizada por mediciones adquiridas tanto

a nivel celular como así también por mediciones adquiridas en la superficie corporal.

2.1.2. Electrocardiograma (ECG)

Un Electrocardiograma (ECG) describe la actividad eléctrica del corazón adquirida por electrodos colocados en la superficie corporal. Las variaciones de voltaje medidas por los electrodos son causadas por los potenciales de acción generados por las células cardíacas cuando las mismas se contraen. El latido resultante en el ECG es manifestado por una serie de ondas, cuyas morfologías y características temporales contienen información que es usada para diagnosticar enfermedades que son causadas por anomalías en la actividad eléctrica del corazón. El patrón temporal que caracteriza la ocurrencia de latidos sucesivos también es muy importante en el diagnóstico de patologías cardíacas. Cuando se iniciaron las primeras mediciones de ECG se realizaban en condiciones favorables para la misma, como puede ser la medición en posición de descanso, pero luego pasó a ser tomado en situaciones no tan favorables como pueden ser ambientes clínicos y ambulatorios donde los algoritmos de procesamiento son esenciales para la extracción de información de los ECG dado a las diversas deformaciones que puede sufrir la medición a causa de distintas fuentes de ruido.

El ECG describe las diferentes fases eléctricas de un ciclo cardíaco y representan la sumatoria en tiempo y espacio de los potenciales de acción cardíacos generados por millones de células cardíacas, por lo que el ECG provee una medición de las corrientes eléctricas generadas en el fluido extracelular generadas por dichos cambios de potencial. Las formas de onda producidas durante la despolarización y la repolarización producen una desviación de una línea base la cual se corresponde con el estado de reposo de las células. Las ondas correspondientes con la despolarización son generalmente más empinadas y más picudas que las ondas correspondientes con la repolarización, que son mas suaves, como se puede observar en la Figura 2.

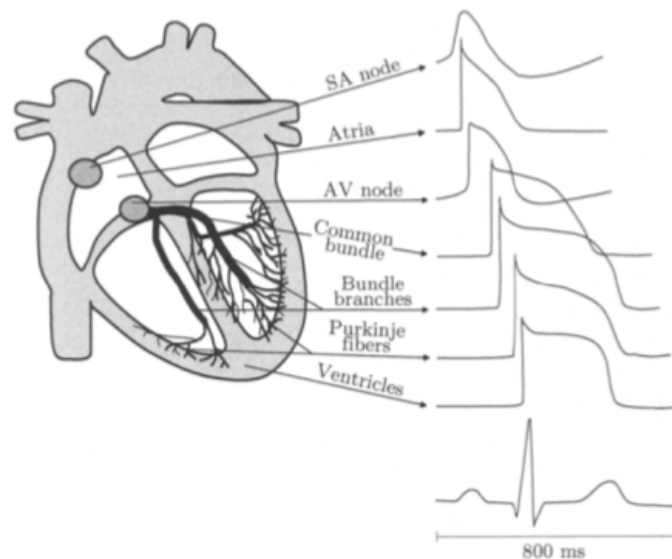


Figura 3: Morfología y tiempos de distintos potenciales de acción localizados en distintas regiones del corazón, y el ciclo cardíaco medido en el ECG.

Para cada instante de un ciclo cardíaco dado, sólo un grupo de células en el miocardio están despolarizadas. Cada grupo de células despolarizadas se puede representar como un dipolo equivalente con su vector asociado, el cual describe su posición, orientación y magnitud variante con el tiempo. Dichos dipolos pueden entonces sumarse para representar un vector “dominante”, el cual describe la dirección principal del impulso eléctrico. Dependiendo de la localización de los electrodos de adquisición del ECG, se pueden obtener diversas visualizaciones de la sumatoria de los potenciales de acción cardíacos, lo que se conocen como distintos canales de ECG. Existen distintas configuraciones de electrodos que se suelen utilizar, donde predominan distintas características de la onda observada.

Los ECG típicamente son registrados con una configuración multicanal, incluyendo canales unipolares, bipolares o ambos. Un canal *unipolar*, refleja la variación en el voltaje en relación a un

electrodo de referencia, el cual está localizado de forma tal que su voltaje se mantenga prácticamente constante. Un canal *bipolar* refleja la variación en el voltage entre dos electrodos, por ejemplo, entre un electrodo colocado en el brazo derecho y otro colocado en el brazo izquierdo.

2.1.3. Características del ECG

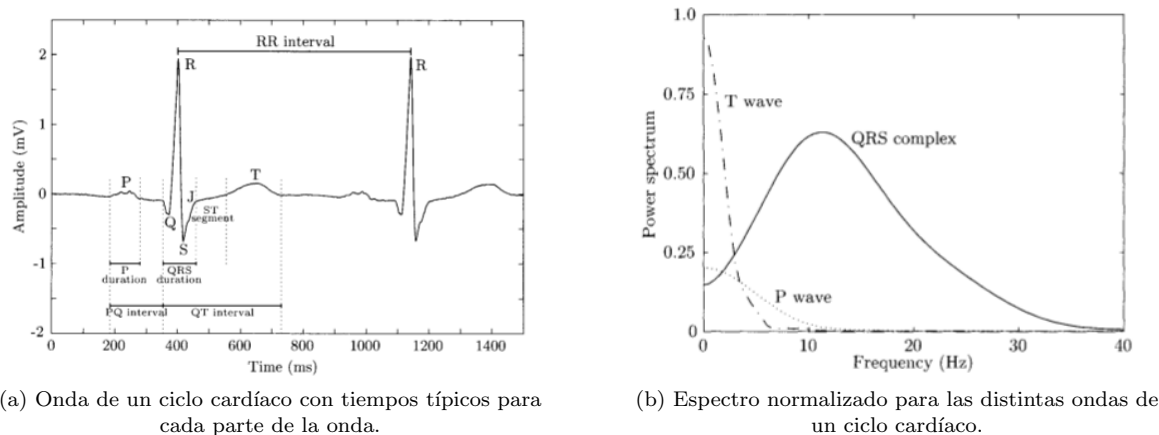


Figura 4: Características temporales y frecuenciales de un ciclo cardíaco

En la Figura 4 se muestra la forma de onda de dos ciclos cardíacos enteros junto con su contenido espectral típico. La *Onda P* refleja la despolarización auricular, y en la mayoría de los canales, tiene una polaridad positiva y una morfología suave y monofásica con una duración típica menor a 120ms . La ausencia de esta onda podría sugerir, por ejemplo, que el latido tuvo origen en los ventrículos. Una *onda P* normal es generalmente considerada de baja frecuencia. El *complejo QRS* refleja la despolarización ventricular, con una duración típica de 70ms a 110ms . La morfología del *complejo QRS* es muy variable y altamente dependiente del origen del latido. Como es la componente de mayor intensidad en el ECG, es la primer onda que se identifica en los algoritmos de análisis. Dichos algoritmos, como el que se presenta en este trabajo, producen los puntos de referencia necesarios para un posterior análisis más profundo de las características del ECG. El *segmento ST* no es en sí una onda, pero representa el intervalo temporal durante el cual los ventrículos se mantienen en un estado de despolarización. La *onda T* refleja la repolarización ventricular y sucede alrededor de 300ms luego del *complejo QRS*, la posición de la misma es altamente dependiente del ritmo cardíaco, comprimiéndose hacia el *complejo QRS* en ritmos cardíacos más altos. El *intervalo RR* representa la longitud temporal de un ciclo cardíaco, medido entre dos *ondas R*, es el ritmo fundamental en cualquier ECG, y es usado para caracterizar diferentes arritmias como así también en estudios de variabilidad del ritmo cardíaco.

2.1.4. Ruido en el ECG

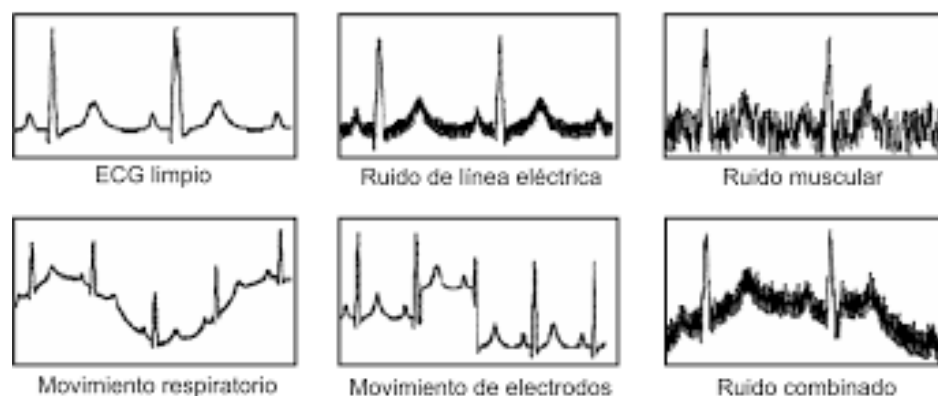


Figura 5: Distintos tipos de ruidos en ECG.

- *Ruido de línea base*: Es una actividad de muy baja frecuencia en el ECG, la cual puede interferir con el análisis de la señal. La misma puede resultar por diversas causas como pueden ser respiración del individuo, movimientos del cuerpo y mal contacto de los electrodos de adquisición. En muchos casos el ruido de línea base sobrepasa el nivel del complejo QRS en gran medida. Su espectro suele considerarse menor a 1Hz pero esta cota puede aumentar en mediciones realizadas con el sujeto realizando algún ejercicio.
- *Movimiento de electrodos*: Es principalmente causado por la piel del sujeto estirándose, lo cual genera cambios en la impedancia de la piel alrededor del electrodo de adquisición. Son más problemáticos que el ruido de línea base, dado que su contenido espectral en general se solapa con todo el espectro del ECG. El ruido a causa de movimiento de electrodos es un problema crítico en el contexto de medición de ECG ambulatorios, donde son la primer causa de latidos detectados erróneamente.
- *Interferencia de ruido de línea*: Es causado por una mala puesta a tierra del equipamiento de medición de ECG y por equipos cercanos generando interferencia. En líneas generales se puede remover mediante un filtrado clásico.
- *Ruido electromiográfico*: Es generado por la actividad eléctrica de los músculos durante períodos de contracción de los mismos. Sus características son muy variables dado que se debe a movimientos propios de los músculos, por lo que pueden contener componentes de alta frecuencia en casos de movimientos repentinos, como así también contenidos más bien estacionarios.
- *Actividad respiratoria*: Generan cambios en la morfología latido a latido, a causa de movimientos del pecho, cambios en la posición del corazón y cambios en la conductividad pulmonar.

Los distintos tipos de ruidos que pueden llegar a afectar en la adquisición del ECG se muestran en la Figura 5. Existen diversas técnicas que se pueden emplear para filtrar algunos de estos tipos de ruido.

2.1.5. Procesamiento de ECG

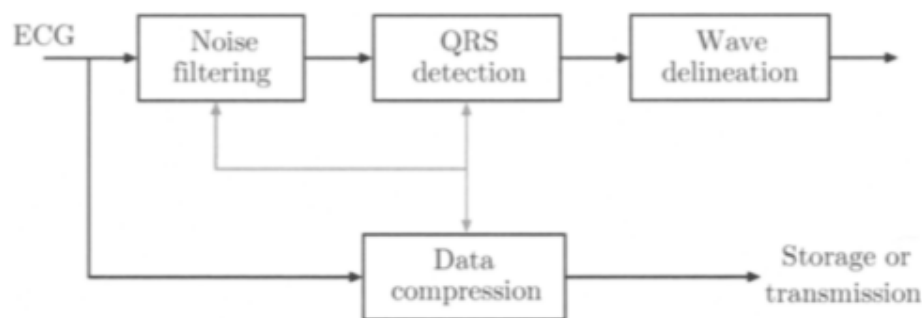


Figura 6: Diagrama clásico de procesamiento de un ECG.

Hasta el momento, ningún sistema ofrece una solución “universal” para el análisis de ECG, sino que cada sistema está diseñado para procesar cierto tipo de ECG registrado bajo ciertas condiciones, por ejemplo, sistemas de interpretación de ECG tomados en situación de descanso, monitoreo de ECG ambulatorio, monitoreo de ECG de terapia intensiva y otros. Todos los sistemas tienen en común un conjunto de algoritmos que acondicionan la señal a trabajar respecto a los distintos tipos de ruido ya mencionados, obtener las mediciones básicas necesarias del ECG, y guardar los resultados de alguna forma eficiente, como se muestra en la Figura 6.

2.1.6. Detección del complejo QRS

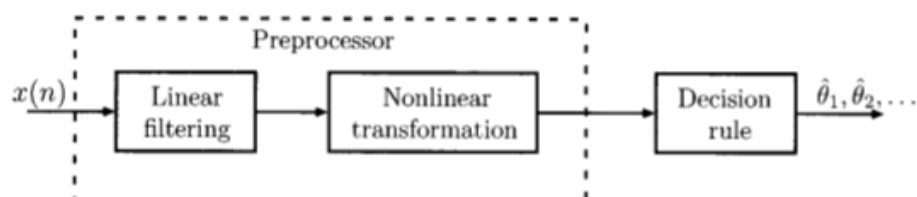


Figura 7: Diagrama en bloques de un detector de complejos QRS típico.

Un detector de complejos QRS debe ser capaz de detectar muy diversas morfologías de los mismos para que pueda ser de utilidad. Asimismo el detector no debe confiar en la característica rítmica de los latidos, sino considerar que los mismos pueden suceder en cualquier momento luego del último latido detectado. La mayoría de los detectores de complejos QRS descriptos en la literatura [3] fueron desarrollados a partir de un razonamiento preciso y experimentación, pero pueden describirse, a grandes rasgos, como el diagrama en bloques presentado en la Figura 7.

Existen detectores de complejos QRS típicamente utilizados y ya descriptos en la literatura, como pueden ser:

- *Pan-Tompkins* [4]: Basado en una serie de filtros pasa altos, una operación no lineal (señal elevada al cuadrado) y un umbral adaptativo.
- *Wavedet* [5]: Basado en la *Transformada Wavelet*. Utiliza un análisis frecuencial para los distintos intervalos temporales del registro.
- *Threshold Independent Algorithm* [6]: Basado en el *Dynamic Plosion Index*. Utiliza un análisis netamente temporal intentando detectar diferencias súbitas del *Plosion Index* en ventanas temporales.
- *Versatile Detector of Pseudo-Periodic Patterns* [7]: Basado en un análisis estadístico de los máximos de la señal luego de aplicar un *Matched Filter* [8].

Si bien estos detectores de complejos QRS tienen sus fortalezas e incluso funcionan bien para muchos ámbitos de ECG, cuando se los utiliza para detección de complejos QRS de otra naturaleza, como pueden ser en ECG de roedores, no tienen un buen rendimiento.

2.2. Redes neuronales

2.2.1. Generalidades

Una *Red Neuronal Artificial* puede ser definida como “Un sistema de computación hecho de un número de elementos simples de procesamiento, altamente interconectados, los cuales procesan información por su respuesta de estado dinámica a impulsos externos” [9] [10]. En gran medida este tipo de modelo computacional fue inspirado por la forma en la que las neuronas del cerebro de los seres humanos se interconectan para procesar información.

2.2.2. Conexiones y analogías biológicas

La unidad básica de computación del cerebro es la *neurona*. El sistema nervioso de los seres humanos está compuesto por alrededor de 86 billones de neuronas, con alrededor de 10^{14} conexiones (*sinápsis*) entre sí.

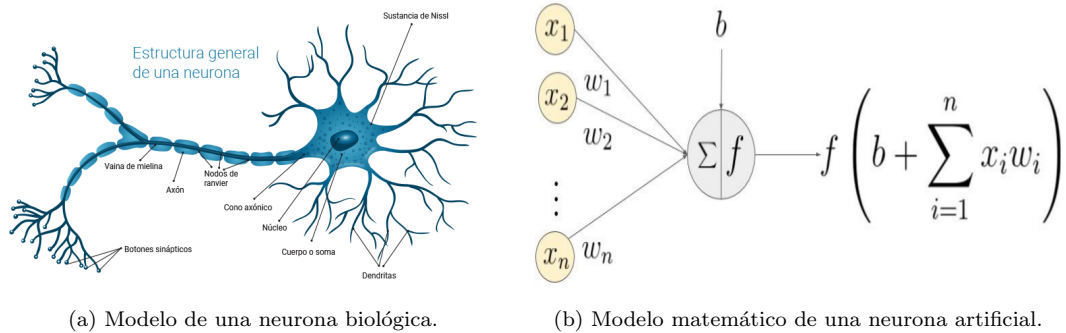


Figura 8: Modelo biológico y modelización matemática de una neurona.

En la Figura 8 se muestra la modelización matemática utilizada para representar una neurona biológica. Análogamente con el cerebro, la unidad básica de computación de una red neuronal también es una *neurona*, pero en este caso utilizamos una modelización matemática. Dicha neurona recibe impulsos de otras neuronas o por eventos externos, y calcula una salida, la cual a su vez puede ser la entrada de otra/s neurona/s y así sucesivamente, conformando la red neuronal.

$$y = f\left(b + \sum_{i=1}^n x_i w_i\right) \quad (1)$$

Tabla 1: Analogía entre la neurona biológica y la modelización matemática

Componente	Neurona Biológica	Neurona Artificial
<i>Entradas</i>	<i>Dendritas</i>	$x_i w_i$
<i>Acumulación</i>	<i>Núcleo</i>	$b + \sum_{i=1}^n x_i w_i$
<i>Umbral</i>	<i>Núcleo</i>	f
<i>Salida</i>	<i>Axión</i>	y

En la Tabla 1 se muestran las analogías aplicables entre las neuronas reales y las artificiales. En las neuronas reales, básicamente las dendritas llevan una señal hacia el núcleo de la neurona, donde

todas se suman. Si la suma de todos los aportes de las dendritas supera un cierto umbral, la neurona puede disparar un pico de señal a través de su axión. Analogamente, en las neuronas artificiales, cada una de las n entradas de una neurona, tiene un *peso* asociado, w_i el cual es multiplicado por cada uno de los valores de entrada x_i , luego se suman todas las entradas ya multiplicadas por su peso, se le suma un *bias* b , y se evalúa el resultado con la *función de activación* f asociada a la neurona, para así calcular la salida y .

2.2.3. Arquitectura de una red neuronal

Si bien las redes neuronales están compuestas por neuronas artificiales, las mismas pueden tener diversos tipos y conexiones entre sí. Se definen a continuación la terminología utilizada a la hora del diseño de los modelos de redes neuronales:

- *Capa neuronal*: Una capa está compuesta por varias neuronas con sus respectivos pesos y bias, compartiendo su función de activación.
- *Capa de entrada*: Está compuesta por los *nodos de entrada*. En esta capa no se realiza ningún procesamiento, solamente sirve a fines de transmitir la información a la próxima capa.
- *Capa/s oculta/s*: En todas las capas ocultas se produce el procesamiento intermedio. Realizan los cálculos y luego transfieren los resultados hacia la próxima capa. Puede haber más de una capa oculta en la red neuronal. Las arquitecturas de redes neuronales que contienen varias capas ocultas, se las denomina *redes neuronales profundas*.
- *Capa de salida*: Es la última capa de la red neuronal. Es la encargada de mapear el resultado final en la configuración adecuada. Por ejemplo, para un clasificador de imágenes que clasifique modelos de autos, tendrá tantas neuronas de salida como modelos de auto pueda clasificar.
- *Regla de aprendizaje*: Es la regla mediante la cual se modifican los parámetros de cada una de las capas de la red neuronal para que las salidas obtenidas se adecúen a las deseadas.

2.2.4. Tipos de capas neuronales básicas

Como ya se mencionó, existen diversos tipos de capas neuronales que se utilizan en el diseño de las redes neuronales. A continuación se explican brevemente los tipos más básicos:

- *Capa lineal*: Cada neurona de la capa anterior está conectada con cada una de las neuronas de esta capa. La cantidad de neuronas que contenga esta capa es un parámetro de la misma.

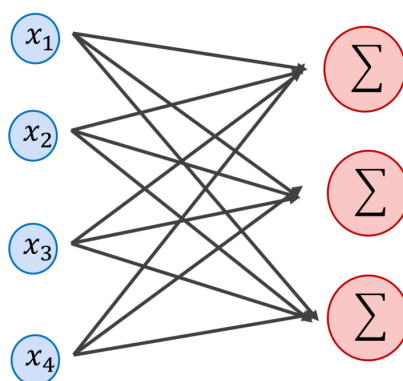


Figura 9: Representación gráfica de una capa lineal.

- *Capa convolucional*: Solo un cierto número de neuronas de la capa anterior está conectada con cada una de las neuronas de esta capa. La cantidad de neuronas de la capa anterior que están conectadas a cada una de las neuronas de esta capa y su distribución son parámetros de la capa.

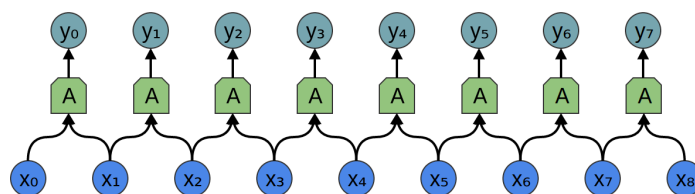


Figura 10: Representación gráfica de una capa convolucional.

- *Capa dropout*: Este tipo de capas son útiles para evitar el *sobreentrenamiento*, explicado en la Sección 4.1.4. En cada período de entrenamiento anulan aleatoriamente una cantidad de neuronas, generando salida nula en dichas neuronas. La cantidad de neuronas a anular es un parámetro de la capa.

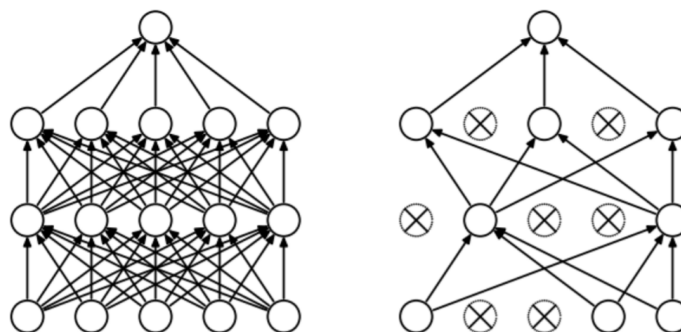
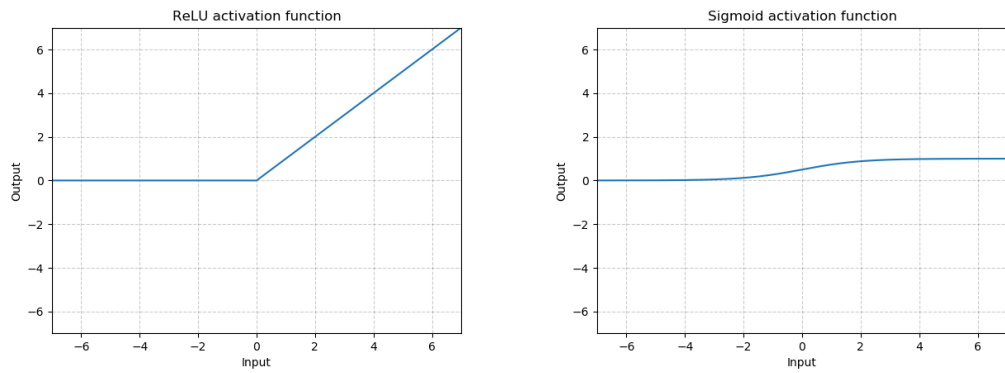


Figura 11: Representación gráfica de tres capas dropout. A la izquierda se observa como serían las mismas capas sin el efecto dropout.

2.2.5. Tipos de funciones de activación básicas

Como ya fue mencionado, la suma de todas las entradas, con sus respectivos pesos afectados, es luego evaluada con una función de activación, para finalmente generar la salida correspondiente. Dichas funciones de activación deben tener ciertas características para ser utilizadas, como ser derivables en todo punto, e idealmente son no lineales. De esta forma, las redes neuronales forman una capacidad muy potente para resolver problemas complejos. También son útiles a la hora de restringir los límites de ciertas entradas o salidas de las distintas capas neuronales, ya que pueden tener imágenes acotadas para un dominio no acotado. A continuación se presentan las funciones de activación más básicas:

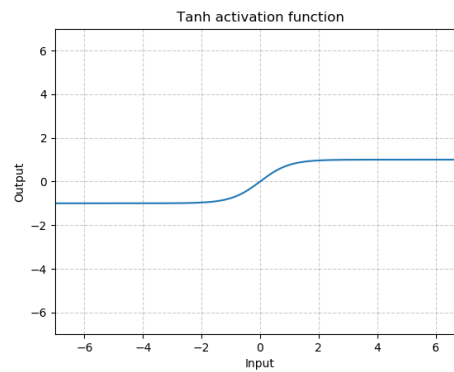


(a) Función de activación ReLU. Se define como:

$$ReLU(x) = \max(x, 0)$$

(b) Función de activación Sigmoide. Se define como:

$$Sigmoide(x) = \frac{1}{1+e^{-x}}$$



(c) Función de activación Tangente hiperbólica. Se define como: $tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

Figura 12: Funciones de activación básicas.

2.2.6. Proceso de aprendizaje de la red neuronal

Una vez definida la arquitectura de la red neuronal en su totalidad, se procede a “entrenar” la misma para que “aprenda” las necesidades de la aplicación en particular. Para este proceso es necesario contar con datos fiables y conocer los resultados deseados para todos los tipos de datos de entrada que se dispongan. Dichos datos son separados en tres sets, llamados sets de *entrenamiento*, *validación* y *evaluación*.

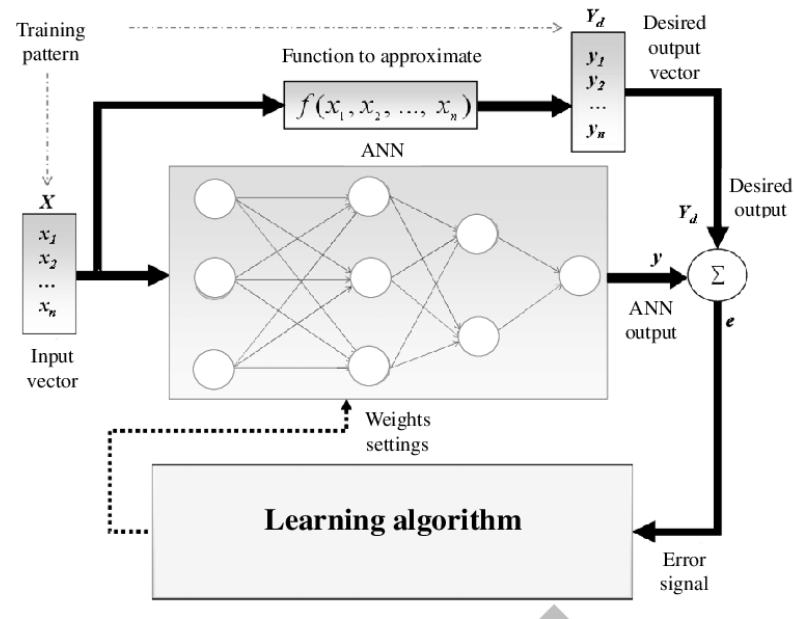


Figura 13: Representación gráfica del proceso de aprendizaje de una red neuronal.

En la Figura 13 se muestra un diagrama representativo del proceso de aprendizaje de una red neuronal. El proceso de aprendizaje de la red neuronal es iterativo. En un inicio, todos los parámetros internos de la red toman un estado inicial aleatorio. Intuitivamente, con estos parámetros no se obtendrán resultados satisfactorios, ya que la red no se ajustó a ningún par entrada/salida hasta el momento. El proceso continúa entonces evaluando iterativamente sucesivas entradas y comparando los resultados obtenidos con los resultados esperados, obteniendo así un error, y pudiendo corregir los parámetros internos de la red a partir de la regla de aprendizaje fijada. Este proceso se repite a lo largo del set de entrenamiento, pudiendo evaluar varias veces cada uno de los datos del set. Dicha cantidad de repeticiones a lo largo del set de entrenamiento es un *hiperparámetro* de la red neuronal, ya que no es parte interna de la red, pero sí es un factor clave en el aprendizaje de la misma.

Una vez dado por terminado el proceso de entrenamiento de la red neuronal, se procede a evaluar la misma. Para dicho cometido se evalúa cada par entrada/salida en el set de validación, el cual no debe contener ningún dato que se encuentre en el set de entrenamiento. En base a herramientas de calificación de algoritmos, se puede saber si los resultados obtenidos están en los niveles esperados o si es necesario corregir la arquitectura de la red neuronal. En caso de obtener resultados no satisfactorios, se corrige la arquitectura de la red neuronal y se vuelve a la etapa de entrenamiento con la arquitectura corregida, para luego volver al proceso de validación, hasta obtener resultados satisfactorios.

Una vez que los resultados en el proceso de validación sean satisfactorios, se procede al proceso final de evaluación. Para dicho proceso se utilizan los datos en el set de evaluación, el cual contiene datos que no deben estar ni en el set de entrenamiento ni en el set de validación. Los resultados obtenidos en éste proceso serán los resultados finales de la red neuronal.

3. Materiales

3.1. Registros de ECG de seres humanos

Para todo el desarrollo pertinente a la detección de complejos QRS en registros pertenecientes a seres humanos fueron utilizadas dos bases de datos disponibles públicamente, explicadas brevemente a continuación:

- *MIT-BIH Arrhythmia Database* [11]: La base de datos MIT-BIH contiene 48 registros de 30 minutos cada uno, obtenidos de 47 individuos distintos en el laboratorio BIH de arritmia entre

1975 y 1978. Los registros fueron digitalizados a una frecuencia de muestreo de 360Hz por canal con una resolución de 11bits en un rango de 10mV . Dos o más cardiólogos anotaron independientemente cada uno de los registros, y luego las discrepancias fueron resueltas. Contiene aproximadamente 110,000 latidos en total.

- *INCART Database* [12]: Esta base de datos contiene 75 registros extraídos de 32 grabaciones realizadas con *Holter*. Los datos fueron provistos por el Instituto de Tecnologías Cardíacas de San Petersburgo. Cada registro es de 30 minutos y contiene los 12 canales estándares, con una frecuencia de muestreo de 257Hz . Contiene aproximadamente 175,000 latidos en total.

Cada registro en cada una de las bases de datos mencionadas, contiene la información acerca de los lugares temporales donde ocurre cada uno de los latidos anotados por profesionales. Esta información será luego la forma de corroborar si la red neuronal está detectando correctamente los latidos.

3.2. Registros de ECG de roedores

Para el desarrollo y evaluación del detector en los datos pertenecientes a ECG de roedores, se utilizaron registros provistos por el grupo de investigación con el que colaboramos en UNCUIYO.

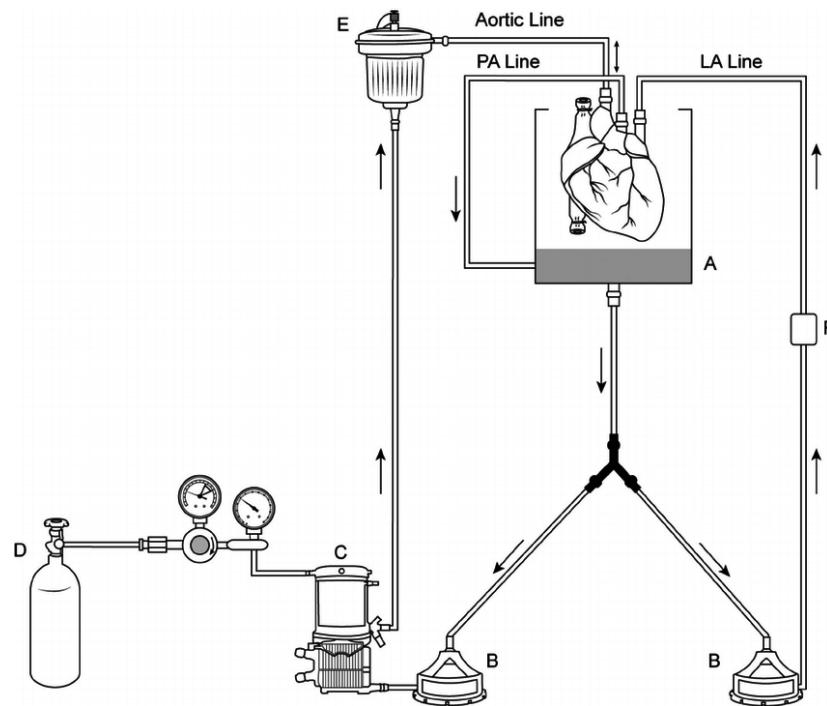


Figura 14: Preparación Langendorff utilizada por el grupo de investigación con el que se colabora en UNCUIYO.

Para obtener los ECG de los roedores, se utiliza una preparación como la que se muestra en la Figura 14 llamada *preparación de Langendorff* [13]. La misma consta de explantar el corazón de un mamífero vivo, colocándolo en una solución similar a la que está inmerso el mismo en el cuerpo vivo, para luego perfundirlo a través de la aorta con una sustancia oxigenada, simulando la sangre oxigenada provista por los pulmones en una condición normal. Esta presión negativa causa que la aorta mantenga cerrada la válvula aórtica, desviando el flujo en su totalidad hacia el orificio de las coronarias nutriendo la masa ventricular, sin existir un llenado ventricular real. De esta forma, el corazón explantado puede seguir latiendo durante un tiempo (algunas horas en general) pudiendo realizar mediciones que no serían posibles en un ser vivo.

Con la *preparación de Langendorff*, proceden a realizar dos mediciones, explicadas a continuación:

- Medición del *Potencial de acción*: Medición directa en una célula muscular, para así medir el potencial de acción transmembrana de la misma.
- Medición del *Campo eléctrico lejano*: Medición que puede asemejarse a una medición de ECG. La diferencia entre las mismas, es que en un ECG tradicional se colocan electrodos directamente en la superficie corporal del individuo del cual se registra el mismo.

En particular, en los registros que nos fueron provistos, ambas mediciones se realizan durante un período de 30 minutos. Los primeros 10 minutos, la solución en la que se encuentra el corazón explantado, se encuentra en condiciones normales. Los próximos 10 minutos, se procede a inducir la solución en hipocalemia es decir, una reducción en el contenido de potasio, lo cual es un contexto que tiende a generar arritmias en el corazón. Los últimos 10 minutos se procede a normalizar el contenido de potasio en la solución. [14]

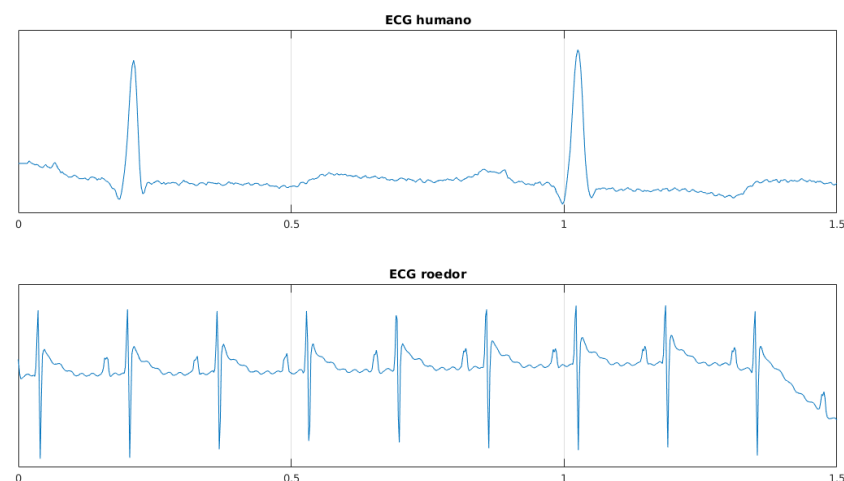


Figura 15: Comparación entre un ECG de un ser humano (arriba) y un ECG de un roedor (abajo).

Cabe destacar que las formas de onda de un ECG perteneciente a un roedor, son iguales a las de un ECG obtenido de un ser humano, pero al tener un corazón más pequeño, los tiempos de cada una de las ondas descritas en la Sección 2.1.3, se “comprimen” como se puede observar en la Figura 15. Esta es una de las principales razones por las cuales los algoritmos estándares, como los descritos en la sección 2.1.6, no obtienen buenos resultados en este tipo de registros.

Al igual que en los registros de ECG de seres humanos, se tiene la información de la ocurrencia temporal de los latidos en cada registro.

3.3. Lenguaje de programación y APIs utilizadas

Para todo el desarrollo del proyecto, se utilizó el lenguaje de programación *Python* [15], ampliamente utilizado en ámbitos como *Data Science* y *Procesamiento de señales*, en conjunto con la librería *PyTorch* [16] utilizada para todo lo pertinente a *redes neuronales*.

4. Metodología

4.1. Preparación de los datos a utilizar

4.1.1. Canales a utilizar de cada registro

Se utilizó únicamente el canal *MLII* en cada uno de los registros a analizar de ambas bases de datos pertenecientes a ECG de seres humanos y únicamente el canal perteneciente al ECG en los datos pertenecientes a ECG de roedores para los tres sets de datos (entrenamiento, evaluación y validación). Los registros que no contuvieran el canal *MLII* en los registros de ECG de humanos fueron descartados. Se tomó esta estrategia para que el objetivo de la red neuronal fuera únicamente aprender ciertas morfologías, y no tener que darle una complejidad más alta aún, considerando que la red neuronal ya tiene que aprender a discernir entre qué es un latido y qué no lo es.

4.1.2. Resampleo, filtrado y ventana temporal utilizada

Una característica de las redes neuronales, es que su capa de entrada tiene un tamaño fijo. Al estar trabajando con señales temporales, esto implica que habrá una relación innata entre la frecuencia de muestreo del registro, y la ventana temporal de análisis que se le presenta a la red neuronal.

En particular, a los ECG de seres humanos se aplicó un filtro pasa altos de muy baja frecuencia ($2Hz$) para eliminar ruidos de línea de base para facilitar el aprendizaje de la red neuronal.

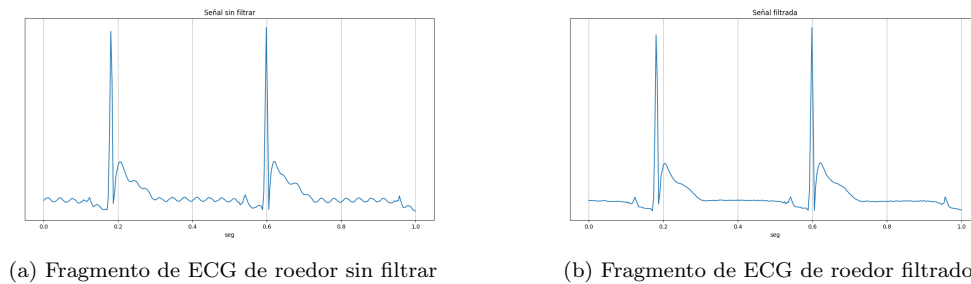


Figura 16: Fragmento de un registro perteneciente a un ECG de roedor antes y después de filtrarlo.

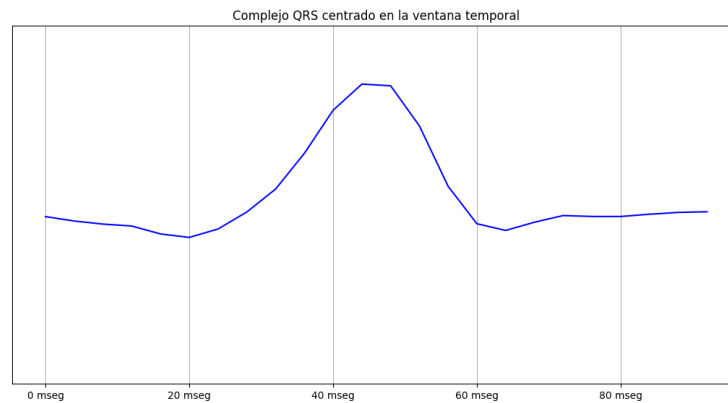
En la Figura 16 se muestra un fragmento de un registro perteneciente a un ECG de roedor antes y después de ser filtrado. Se puede observar, en la señal sin filtrar, una alta componente de ruido de $50Hz$, causada por las condiciones de medición de los registros. Por esta razón, en los registros pertenecientes a ECG de roedores se utilizó el mismo filtro pasa altos con frecuencia de corte en $2Hz$, pero también un *filtro notch* con frecuencia central en $50Hz$ para disminuir un poco el ruido de línea, como se puede observar en la señal ya filtrada.

Cabe destacar que para el filtrado de las señales se utilizó un *filtrado de fase cero*, ya que se desea mantener las características de envolvente y de forma de la señal lo mayor posible.

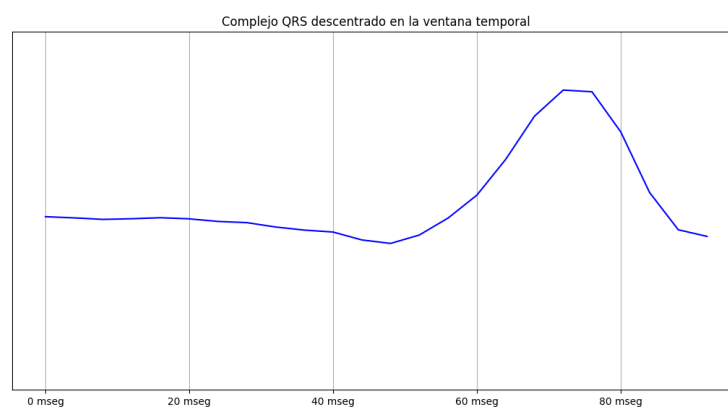
4.1.3. Aumentación de datos

Como fue explicado en la Sección 2.2.6, la red neuronal “aprenderá” a discernir qué es un latido de qué no lo es, a partir de un proceso iterativo.

Para que la red neuronal pueda obtener una buena eficiencia, se debe considerar el hecho de que las ventanas temporales utilizadas para tomar las distintas secciones de cada registro, pueden contener o no contener un latido, no obstante, el latido puede no encontrarse en el centro de la ventana temporal, como se muestra en la Figura 17. Cada ventana temporal generada en cada registro debe



(a) Complejo QRS centrado en la ventana temporal.



(b) Complejo QRS descentrado en la ventana temporal.

Figura 17: Ubicaciones distintas del complejo QRS en una ventana temporal.

ser clasificada como *latido* o *no latido*, para luego comparar el resultado obtenido a través de la red neuronal, con el resultado verdadero. Esto implica que se debe considerar como *latido* a todas las ventanas temporales que contengan una parte o la totalidad de un complejo QRS. A este proceso de generar múltiples datos a partir de transformaciones, en este caso un desplazamiento temporal, se lo conoce como *data augmentation* o *augmentación de datos*. En la preparación de los datos, tanto en los registros de ECG de seres humanos como en los registros de ECG de roedores, se aplicó dicha técnica para contemplar todas las ubicaciones donde puede ubicarse el complejo QRS en la ventana temporal.

4.1.4. Sobreentrenamiento

Como fue mencionado en la Sección 2.2.3, las redes neuronales están conformadas por múltiples capas neuronales, cada una con una cantidad de parámetros que irán siendo actualizados en el proceso de entrenamiento. Un problema usual en este tipo de algoritmos, es el *overfitting* o *sobreentrenamiento*.

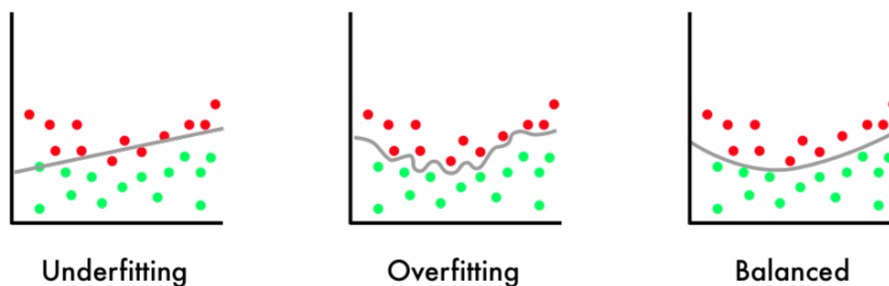


Figura 18: Visualización del sobreentrenamiento.

Izquierda: Predicción que le falta generalizar más.

Centro: Predicción sobreentrenada.

Derecha: Predicción balanceada (objetivo a lograr).

Se denomina *overfitting* al fenómeno en donde un clasificador se especializa tanto en las muestras correspondientes al set de entrenamiento, que no puede generalizar bien en muestras distintas a las del set mencionado. Este fenómeno se puede dar por diversas razones, siendo la principal darle demasiada *capacidad* a la red neuronal, es decir, que la red neuronal tenga tantos parámetros para aprender que termina adaptándose por demás a los datos de entrenamiento. Otra de las causas de este fenómeno es entrenar la red durante un tiempo excesivo, donde los parámetros tal vez no sean excesivos, pero se terminan sobreajustando a las muestras en el set de entrenamiento, de forma tal que pierden poder de generalización.

Si bien al hacer *aumentación de datos* obtenemos muchos más datos de los que inicialmente se contaba, al conocer las señales a trabajar, se sabe que van a haber muchas más secciones en un ECG de seres humanos donde no habrán latidos, que zonas donde habrán latidos. Esto podría llegar a generar un sobreentrenamiento de un tipo distinto a los mencionados en el párrafo anterior, ya que la red neuronal aprendería con mucha más exactitud a determinar si en una ventana temporal no hay latido en comparación con su exactitud para determinar si en una ventana temporal hay un latido. Para evitar esto, se determinó experimentalmente una relación de *nolatido : latido* aproximadamente 3 : 1 tomando a lo largo de una vasta cantidad de registros secciones temporales de igual tamaño y determinando si en la misma había latido o no. En base a esta relación, una vez realizada la *aumentación de datos* en cada uno de los registros sobre los cuales armar los sets de entrenamiento, evaluación y validación, se procedió a tomar al azar de forma no repetitiva tres no latidos por cada latido, para así finalmente conformar los tres sets.

4.1.5. Separación de datos para cada set

El total de los datos fue separados en los sets de entrenamiento, validación y evaluación como fue explicado en la Sección 2.2.6. El porcentaje del total de los datos con el cual se construye cada uno de dichos sets, depende en gran medida de la cantidad de datos disponibles, así como en la variabilidad de los datos. Si la variabilidad es chica, se pueden dar porcentajes más pequeños a los sets de validación y evaluación, ya que la muestra que se tome, más allá de ser pequeña, representará en gran medida a la totalidad de los datos.

Tabla 2: Porcentaje del total de los datos destinados para los sets de entrenamiento, validación y evaluación.

Set	Porcentaje
<i>Entrenamiento</i>	50 %
<i>Validación</i>	25 %
<i>Evaluación</i>	25 %

Para ambos casos de estudio a lo largo de este proyecto, se tomaron los porcentajes indicados en la Tabla 2 respecto del total de los datos disponibles.

4.2. Arquitecturas neuronales planteadas

4.2.1. Primer arquitectura

En una primer instancia se utilizó una arquitectura neuronal como la descripta en la Tabla 3. Se la diseñó tomando como base ejemplos encontrados en la documentación de la API utilizada [16].

Tabla 3: Primer arquitectura neuronal planteada.

Número de capa	Tipo de capa	Parámetros	Tamaño de entrada	Tamaño de salida
1	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	24×1	25×1
2	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	25×1	26×1
3	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	26×1	27×1
4	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	27×1	28×1
5	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	28×1	29×1
6	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	29×1	30×1
7	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	30×1	31×1
8	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	31×1	32×1
9	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	32×1	33×1
10	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	33×1	34×1
11	<i>Lineal</i>	—	34×1	24×1
12	<i>ReLU</i>	—	24×1	24×1
13	<i>Dropout</i>	$p = 0,5$	24×1	24×1
14	<i>Lineal</i>	—	24×1	12×1
15	<i>ReLU</i>	—	12×1	12×1
16	<i>Dropout</i>	$p = 0,2$	12×1	12×1
17	<i>Lineal</i>	—	12×1	1×1
18	<i>Sigmoide</i>	—	1×1	1×1

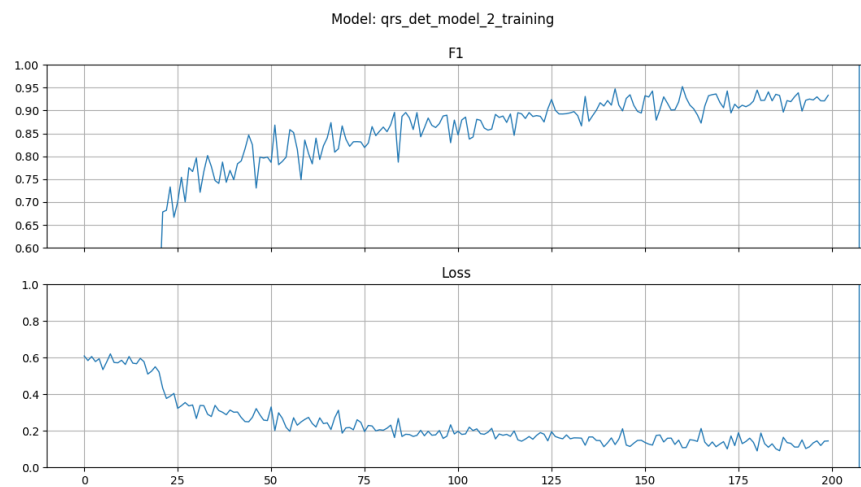


Figura 19: Fragmento del entrenamiento de la primer arquitectura planteada.

Los resultados obtenidos con la arquitectura no fueron lo suficientemente satisfactorios, los mismos se encuentran expuestos en la Sección 5. Asimismo, como se puede observar en la Figura 19, el entrenamiento de la red se ve “errático” o “ruidoso”. Se concluyó en que esto fue debido a que la red neuronal se estaba sobreentrenando, fenómeno explicado en la Sección 4.1.4, por un exceso en su capacidad, dado que para las distintas correcciones efectuadas en cada paso de aprendizaje de la misma varía demasiado el resultado obtenido, y la hipótesis es que en cada paso de aprendizaje la red neuronal debería acercarse de alguna forma un poco más asintótica hacia el resultado ideal, como sucedió con la arquitectura final expuesta en la Sección 4.2.3. En base a la conclusión mencionada, se pasó a la segunda arquitectura, que resultaría en la arquitectura final utilizada en este proyecto.

4.2.2. Modificaciones realizadas a la arquitectura planteada

El criterio adoptado para las distintas modificaciones realizadas a la arquitectura de la red neuronal planteada inicialmente, fue basado en la evolución en el entrenamiento de los distintos intentos de arquitecturas planteados, pero tuvo una naturaleza *iterativa*.

En particular, se inició por convertir las múltiples capas convolucionales planteadas en la primer arquitectura, en una cantidad menor de capas convolucionales, pero con mayor cantidad de canales de salida, entendiéndose que de esta forma las mismas podrían detectar distintas características o *features* de la señal a analizar, obteniendo resultados mejores.

Otra estrategia que se aplicó posteriormente, fue probar distintas funciones de activación que llevaron a resultados pocos satisfactorios, dado que en la mayoría de los casos la red neuronal terminaba en un estancamiento de su entrenamiento con resultados muy pobres, o entrenamientos que se volvían cada vez más erráticos, por lo que se optó por continuar con las funciones de activación tipo *ReLU*.

La última estrategia aplicada, la cual llevó a la arquitectura final y a los mejores resultados, fue optar por confluir en una única capa lineal a partir de la salida de la última capa convolucional, la cual se encarga de mezclar en cierta forma todas las características que las capas anteriores ya analizaron, y además eliminar las capas de *dropout*, dado que la red no parecía sufrir un sobreentrenamiento, probablemente dada su poca profundidad y por ende baja capacidad.

4.2.3. Arquitectura final

Las distintas capas neuronales de la arquitectura final utilizada se encuentra descripta en la Tabla 4. Nótese que la profundidad de la red neuronal final es mucho menor en comparación con la primer arquitectura planteada, sin embargo obtuvo resultados más satisfactorios. Los resultados para las dos bases de datos están expuestos en la Sección 5.

Tabla 4: Arquitectura neuronal final planteada.

Número de capa	Tipo de capa	Parámetros	Tamaño de entrada	Tamaño de salida
1	<i>Convolutacional</i>	$kernel = 8$ $padding = 4$	24×1	25×3
2	<i>ReLU</i>	—	25×3	25×3
3	<i>Convolutacional</i>	$kernel = 6$ $padding = 3$	25×3	26×6
4	<i>ReLU</i>	—	26×6	26×6
5	<i>Convolutacional</i>	$kernel = 4$ $padding = 2$	26×6	27×1
6	<i>ReLU</i>	—	27×1	27×1
7	<i>Lineal</i>	—	27×1	1×1
8	<i>Sigmoide</i>	—	1×1	1×1

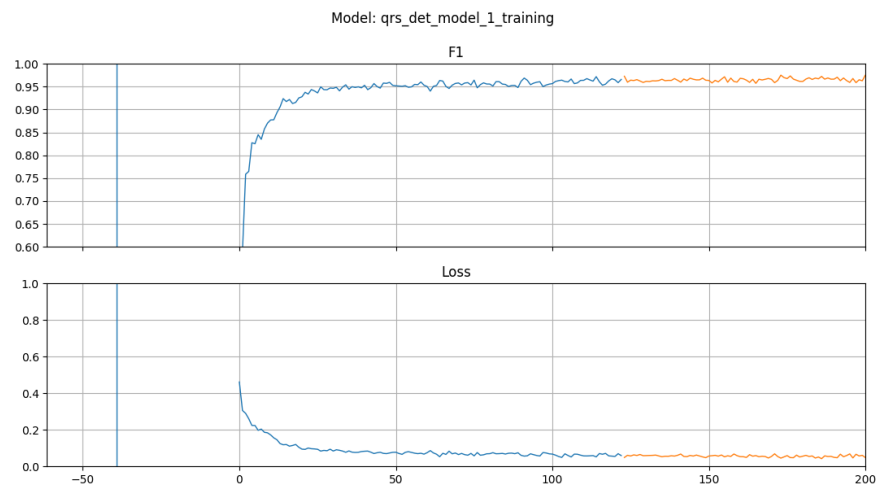


Figura 20: Fragmento del entrenamiento de la arquitectura final planteada.

Como se puede observar en la Figura 20, el entrenamiento de la red neuronal con la arquitectura final utilizada, tiene una evolución mucho menos “ruidosa” que la mostrada en la Figura 19 perteneciente al entrenamiento de la primera arquitectura planteada. Esto parece indicar un entrenamiento mejor encausado a solucionar el problema, como queda demostrado con los resultados.

4.3. Método de evaluación del detector

4.3.1. Matriz de confusión

Para evaluar el detector se utilizó la *matriz de confusión* [17], también llamada *matriz de error*. Es típicamente utilizada para calificar algoritmos de detección, sean multi-clase o uni-clase.

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Figura 21: Matriz de confusión.

La misma consta de generar una matriz como la que se ilustra en la Figura 21, donde se clasifican cada una de las detecciones realizadas por el detector bajo evaluación, donde cada una de las detecciones se puede clasificar como:

- *Verdadero Positivo (TP)*: El detector clasificó la muestra como un **latido** y realmente era un **latido**.
- *Verdadero Negativo (TN)*: El detector clasificó la muestra como un **no latido** y realmente era un **no latido**.
- *Falso Positivo (FP)*: El detector clasificó la muestra como un **latido** y realmente era un **no latido**.

- *Falso Negativo (FN)*: El detector clasificó la muestra como un **no latido** y realmente era un **latido**.

A partir de la *matriz de confusión*, se pueden obtener indicadores representativos de distintos aspectos del algoritmo clasificador, en particular, se utilizaron los siguientes:

$$F1 = \frac{2TP}{2TP + FN + FP}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$MCC = \left(\frac{TPTN - FPFN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} + 1 \right) / 2$$

Donde:

- *F1 Score (F1)*: Es una medida muy utilizada en clasificaciones binarios. Tiene en cuenta únicamente los *verdaderos positivos* sin tomar así en cuenta los *verdaderos negativos*. En situaciones donde es crítico que el clasificador no pierda observaciones positivas, es un indicador muy importante. Sus valores están acotados entre 0 y 1, 0 indicando la ausencia total de detecciones verdaderas correctas, y 1 que la totalidad de las detecciones verdaderas fueron correctas, y que no se omitió ninguna detección negativa.
- *Accuracy o Exactitud (ACC)*: Es una medida un poco más general que el *F1 Score*, ya que tiene en cuenta tanto *verdaderos positivos* como *verdaderos negativos*. Sus valores van entre 0 y 1, 0 indicando que el clasificador no es para nada exacto, y 1 indicando que es totalmente exacto.
- *Coficiente de Correlación de Matthews (MCC)*: Al igual que la *Exactitud* tiene en cuenta tanto *verdaderos positivos* como *verdaderos negativos*, y suele ser una medida más balanceada para evaluar clasificadores binarios. Cabe destacar que la ecuación original fue modificada para que su imagen esté acotada entre 0 y 1, y no entre -1 y 1, a fines de poder compararlo con los otros dos parámetros utilizados.

Cabe destacar que los indicadores expuestos fueron utilizados en las tres etapas de desarrollo de la red neuronal, pero solo se exponen los resultados correspondientes al set de evaluación, en particular, de la arquitectura final utilizada, en la próxima sección, mientras que en la etapa de entrenamiento, son utilizados para detectar un sobreentrenamiento temprano, o un estancamiento en el aprendizaje de la red, y en el set de validación son utilizados para corroborar que la red no se sobreadaptó al set de entrenamiento.

5. Resultados

En esta sección se exponen los resultados obtenidos en el set de evaluación correspondientes a las bases de datos de ECG de seres humanos y ECG de roedores.

5.1. Resultados en ECG de seres humanos con ambas arquitecturas sobre los sets de validación

Tabla 5: Resultados del set de validación en los ECG de seres humanos para ambas arquitecturas.

Arquitectura	Registro	F1 Score	ACC	MCC
Primera	<i>mediana</i>	95,75 %	97,89 %	97,19 %
Primera	<i>promedio</i>	94,23 %	97,09 %	96,27 %
Final	<i>mediana</i>	99,04 %	99,51 %	99,36 %
Final	<i>promedio</i>	97,88 %	98,88 %	98,59 %

En la Tabla 5 se muestran los resultados obtenidos para el mismo set de validación para las dos arquitecturas planteadas a lo largo del trabajo. Si bien los resultados obtenidos en la primera arquitectura pareciesen muy cercanos a los obtenidos en la segunda, en terminos comparativos contra otros algoritmos clásicos que realizan la misma tarea, es una distancia considerable. Además, el sobreentrenamiento mencionado en la Sección 4.2.1 es un problema que no se refleja únicamente en los resultados, por lo cual es útil el análisis realizado respecto al desarrollo del aprendizaje de la red neuronal.

5.2. Resultados en ECG de seres humanos con la arquitectura final en el set de evaluación

Tabla 6: Resultados del set de evaluación en los ECG de seres humanos.

Registro	F1 Score	ACC	MCC
<i>INCART_I04</i>	99,98 %	99,99 %	99,99 %
<i>INCART_I05</i>	97,57 %	98,57 %	98,31 %
<i>INCART_I14</i>	99,86 %	99,93 %	99,91 %
<i>INCART_I21</i>	95,97 %	97,92 %	97,32 %
<i>INCART_I22</i>	99,16 %	99,58 %	99,44 %
<i>INCART_I25</i>	97,80 %	98,90 %	98,53 %
<i>INCART_I31</i>	99,98 %	99,99 %	99,99 %
<i>INCART_I37</i>	99,48 %	99,73 %	99,65 %
<i>INCART_I38</i>	90,92 %	94,46 %	93,53 %
<i>INCART_I40</i>	99,93 %	99,97 %	99,96 %
<i>INCART_I43</i>	99,84 %	99,92 %	99,89 %
<i>INCART_I44</i>	93,99 %	97,06 %	96,04 %
<i>INCART_I47</i>	98,13 %	99,05 %	98,76 %
<i>INCART_I58</i>	74,30 %	89,77 %	86,05 %
<i>INCART_I60</i>	99,37 %	99,68 %	99,58 %
<i>INCART_I65</i>	98,95 %	99,47 %	99,30 %
<i>INCART_I66</i>	96,25 %	98,05 %	97,53 %
<i>INCART_I75</i>	98,75 %	99,37 %	99,17 %
<i>mitdb_107</i>	66,74 %	87,46 %	82,62 %
<i>mitdb_112</i>	97,84 %	98,78 %	98,50 %
<i>mitdb_116</i>	98,04 %	98,92 %	98,66 %
<i>mitdb_122</i>	99,99 %	100,00 %	100,00 %
<i>mitdb_123</i>	99,70 %	99,85 %	99,80 %
<i>mitdb_124</i>	99,80 %	99,90 %	99,87 %
<i>mitdb_200</i>	99,90 %	99,95 %	99,93 %
<i>mitdb_203</i>	99,84 %	99,74 %	99,65 %
<i>mitdb_209</i>	99,24 %	99,62 %	99,49 %
<i>mitdb_213</i>	99,33 %	99,67 %	99,56 %
<i>mitdb_220</i>	90,04 %	99,52 %	99,36 %
<i>mediana</i>	99,16 %	99,58 %	99,44 %
<i>promedio</i>	96,50 %	98,44 %	97,94 %

En la Tabla 6 se disponen los resultados obtenidos para cada registro en el set de ECG perteneciente a seres humanos. Nótese que la red neuronal logró generalizar bastante bien, ya que el rendimiento general baja del 90 % únicamente para dos registros.

5.2.1. Resultados en ECG de roedores

Tabla 7: Resultados del set de evaluación en los ECG de roedores.

Registro	F1 Score	ACC	MCC
1	92,84 %	96,66 %	95,54 %
2	94,68 %	97,43 %	96,55 %
3	50,78 %	83,50 %	76,40 %
4	98,80 %	99,41 %	99,21 %
<i>mediana</i>	93,76 %	97,05 %	96,04 %
<i>promedio</i>	84,28 %	94,25 %	91,92 %

En la Tabla 7 se disponen los resultados obtenidos para cada registro en el set de ECG perteneciente a roedores. Cabe destacar que si bien el rendimiento en este caso es menor al correspondiente

en ECG perteneciente a seres humanos, es un resultado muy positivo.

5.3. Resultados registros sintéticos

Tabla 8: Resultados del set de evaluación de los registros sintéticos generados a partir de patrones obtenidos de los registros pertenecientes a ECG de roedores.

Registro	F1 Score	ACC	MCC
1	100 %	100 %	100 %
2	100 %	100 %	100 %
3	100 %	100 %	100 %
4	100 %	100 %	100 %
<i>mediana</i>	100 %	100 %	100 %
<i>promedio</i>	100 %	100 %	100 %

Bajo la hipótesis de que el ruido excesivo en las señales era un problema, se diseñaron señales sintéticas a partir de recortar varios patrones en los registros disponibles, y luego generando registros sintéticos repitiendo aleatoriamente los distintos patrones a distintos intervalos temporales, y sumándole finalmente un nivel de ruido blanco. Los resultados en estos registros se exponen en la Tabla 8. Como se puede observar, la red neuronal se adaptó completamente al set de registros, indicando que el ruido presente en los registros pertenecientes a ECG de roedores, es sumamente influyente en la detección.

5.4. Comparación con otro algoritmo de detección utilizado actualmente

Tabla 9: Comparación en el set perteneciente a roedores, del rendimiento perteneciente a la red neuronal en contraposición de otro algoritmo públicamente disponible.

Algoritmo	F1 Score (mediana)	F1 Score (promedio)
<i>wavedet</i>	53,48 %	65,01 %
<i>red neuronal</i>	95,68 %	84,28 %

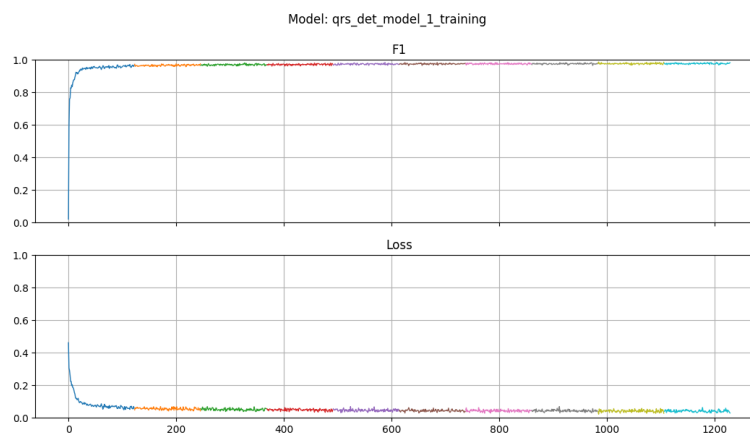
5.5. Proceso de entrenamiento para la red neuronal

En la Figura 22 se muestran los progresos en el entrenamiento de la red neuronal para los datos de ECG de seres humanos y para los ECG de roedores. Nótese que para los ECG de seres humanos, se tiene una variabilidad menor que para los ECG de roedores. Dicho efecto es causado por la alta variabilidad en los registros de ECG de roedores, en donde la red neuronal tiene una dificultad mayor para obtener una buena generalización de las características de los datos.

6. Conclusiones

En el trabajo realizado se diseñó una red neuronal que logra detectar complejos QRS con una fiabilidad alta comparable con otros algoritmos de detección utilizados hoy en día, entendiéndose por fiabilidad alta valores medianos mayores al 99 %, no sólo detectando complejos QRS en ECG pertenecientes a seres humanos, sino también en ECG pertenecientes a mediciones de campo lejano de corazones de roedores explantados, únicamente reentrenando la red neuronal con un set distinto de datos.

Para obtener los calificadores del algoritmo de detección, se lo evaluó en 1801032 *latidos* y 5289786 *no latidos* para los registros pertenecientes a ECG de seres humanos, y se lo evaluó en



(a) Evolución en el entrenamiento para los ECG de seres humanos.



(b) Evolución en el entrenamiento para los ECG de roedores.

Figura 22: Evolución en los entrenamientos de las redes neuronales. Los distintos colores en cada gráfico implican distintas *épocas* de entrenamiento.

79950 *latidos* y 239850 *no latidos*, logrando un $MCC = 99,36\%$ en los ECG de seres humanos y un $MCC = 96,04\%$ en los ECG de roedores.

La red neuronal obtuvo resultados que son comparables en los ECG de seres humanos con otros algoritmos actuales de detección de complejos QRS, y superó ampliamente a otro detector en los ECG de roedores, únicamente reentrenando la red con los datos correspondientes. Cabe destacar que la red no conoce absolutamente ninguna característica acerca de los patrones que tiene que detectar, y sin embargo “aprende” a detectar los mismos con una alta efectividad, lo cual no es un detalle menor, ya que implica una gran potencialidad para solucionar problemas complejos de los cuales no se tengan una solución fácilmente alcanzable.

Como se mencionó en la Sección 5.3, el alto contenido de ruido en las señales de ECG pertenecientes a roedores pareciese afectar mucho a la correcta detección de la red neuronal, lo cual quedó demostrado en parte con los registros sintéticos generados. Una próxima etapa en la colaboración con UNCuyo sería mejorar los elementos de adquisición de datos y las condiciones de la propia adquisición de datos para mejorar la calidad de las señales obtenidas. Dicho efecto también se puede observar en la Sección 5.5 con el progreso de entrenamiento de la red neuronal para los distintos tipos de datos, marcando una alta variabilidad en los datos de los ECG de roedores, a causa del gran contenido de ruido de los mismos.

Los resultados que obtuvo la red neuronal tanto en ECG de seres humanos como en ECG de roedores, es altamente satisfactoria, y en particular para los ECG de roedores, resultaron más que satisfactorias para los colegas de UNCuyo, a quienes se les comentaron los resultados obtenidos, obteniendo una retroalimentación positiva respecto al trabajo realizado.

Referencias

- [1] World Health Organization. (2017) Cardiovascular diseases. World Health Organization. [Online]. Available: http://www.who.int/cardiovascular_diseases/en/
- [2] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier, 2005.
- [3] M. Elgendi *et al.*, “Revisiting qrs detection methodologies for portable, wearable, battery-operated, and wireless ecg systems,” *PLOS ONE*, vol. 9, no. 1, pp. 1–18, 01 2014. [Online]. Available: <https://doi.org/10.1371/journal.pone.0084018>
- [4] J. Pan and W. J. Tompkins, “A real-time qrs detection algorithm,” *IEEE Transactions on Biomedical Engineering*, pp. 230–236, 1985. [Online]. Available: <https://doi.org/10.1109/TBME.1985.325532>
- [5] J. P. Martínez, R. Almeida, S. Olmos, A. Rocha, and P. Laguna, “A wavelet-based ECG delineator: Evaluation on standard databases,” *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 570–581, 2004. [Online]. Available: <https://doi.org/10.1109/TBME.2003.821031>
- [6] A. G. Ramakrishnan, A. P. Prathosh, and T. V. Ananthapadmanabha, “Threshold-independent qrs detection using the dynamic plosion index,” *IEEE Signal Processing Letters*, vol. 21, no. 5, pp. 554–558, May 2014. [Online]. Available: <https://www.doi.org/10.1109/LSP.2014.2308591>
- [7] A. Santini *et al.*, “Versatile detector of pseudo periodic patterns,” *CinC*, 2018. [Online]. Available: <https://doi.org/10.22489/CinC.2018.379>
- [8] G. Turin, “An introduction to matched filters,” *IEEE Transactions on Information Theory*, pp. 311–329, 1960. [Online]. Available: <https://doi.org/10.1109/TIT.1960.1057571>
- [9] D. Fumo. (2017) A gentle introduction to neural networks series. [Online]. Available: <https://towardsdatascience.com/a-gentle-introduction-to-neural-networks-series-part-1-2b90b87795bc>
- [10] C. Olah. (2014) Neural networks, manifolds, and topology. [Online]. Available: <https://colah.github.io/posts/2014-03-NN-Manifolds-Topology/>
- [11] Golderberg *et al.*, “The impact of the mit-bih arrhythmia database,” 2005. [Online]. Available: <https://doi.org/10.13026/C2F305>
- [12] G. Moody and R. Mark, “St petersburg incart 12-lead arrhythmia database,” 2008. [Online]. Available: <https://doi.org/10.13026/C2V88N>
- [13] D. Riascos, E. Baltaxe, and G. Pascual, “La preparación de langendorff: corazón de mamífero aislado y perfundido,” *Universitas Medica*, vol. 45, p. 111, 01 2004.
- [14] E. R. Diez *et al.*, “Melatonin, given at the time of reperfusion, prevents ventricular arrhythmias in isolated hearts from fructose fed rats and spontaneously hypertensive rats,” *Journal of Pineal Research*, vol. 55, no. 2, pp. 166–173.
- [15] The python programming language. [Online]. Available: <https://www.python.org/>
- [16] Pytorch api. [Online]. Available: <https://www.pytorch.org/>
- [17] T. Fawcett, “An introduction to roc analysis,” 2005. [Online]. Available: <https://doi.org/10.1016/j.patrec.2005.10.010>