

Proyecto 3: Procesamiento de señales ECG mediante diferentes filtros.

Luisa F. Enciso¹ y Jader S. Chingal¹.

¹Universidad De Antioquia Facultad De Ingeniería, Bioingeniería, Bioseñales

luisa.enciso@udea.edu.co
jstalyn.chingal@udea.edu.co

I. CONTEXTUALIZACIÓN

Este proyecto tiene como objetivo implementar un flujo de procesamiento para señales de electrocardiograma (ECG), que permita la extracción de características espectrales y la reducción de ruido para su posterior análisis.

Se proponen tres tipos de flujo de procesamiento que pueden incluir varios pasos críticos, como el filtrado pasa-altas, filtrado wavelet modificado y filtrado pasabajas, que permiten limpiar la señal ECG de contaminantes y artefactos. Estos filtros son esenciales para destacar las características relevantes de la señal y facilitar el análisis posterior. Se evalúan cada uno de ellos para determinar el más adecuado. Adicionalmente, el análisis se centra en dos tipos de arritmias: bradicardia sinusal (SB) y fibrilación auricular (AFIB). Para cada señal procesada, se extraerá la frecuencia de máxima potencia utilizando el método de Welch, y se almacenarán los resultados en un dataframe junto con el tipo de patología.

Finalmente, el proyecto también incluirá una comparación entre los resultados obtenidos del flujo de procesamiento propuesto y aquellos del proyecto anterior (señales prefiltradas), usando estadísticas descriptivas y pruebas de hipótesis para evaluar si hay diferencias significativas en los tipos de señales procesadas. Este análisis tiene como fin mejorar la capacidad de diagnóstico de las tecnologías basadas en ECG.

II. MÉTODOS DE REDUCCIÓN DE RUIDO.

A. Robust LOESS (Locally Estimated Scatterplot Smoothing)

Este método se utiliza para eliminar la distorsión o fluctuación lenta de baja frecuencia (0,5 Hz) en la señal ECG, que causa que la línea de base se desplace hacia arriba o hacia abajo de forma irregular. Este problema se presenta principalmente por la respiración del paciente.

El método es una técnica de suavizado de datos que ajusta curvas locales para eliminar las fluctuaciones no deseadas, primero divide los datos en pequeñas ventanas o segmentos. Para cada punto de datos en la señal, considera solo datos vecinos cercanos en lugar de usar toda la serie de datos. Esto

permite crear un ajuste local, luego en cada ventana realiza una regresión ponderada. Esto significa que ajusta una curva a los datos del segmento de manera iterativa para manejar valores atípicos, el método reduce la influencia de estos puntos atípicos hasta que logra un ajuste estable y confiable para la línea de base. Una vez que el ajuste a convergido, el resultado es una curva suavizada que representa la tendencia general de los datos en cada ventana sin la influencia de variaciones bruscas.[1]

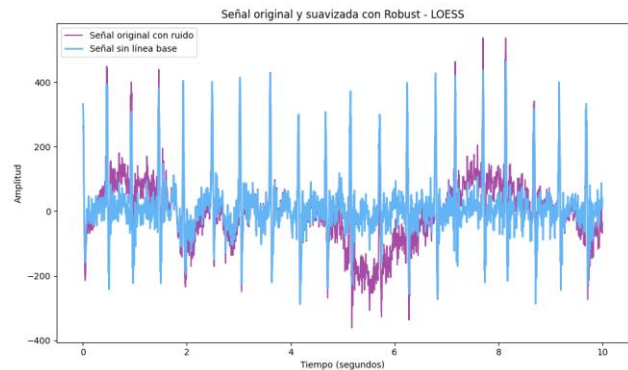


Figura 1: Señal ECG aplicando el método Robust LOESS

Se extrae la señal de un paciente de la base de datos y se le aplica el método Robust LOESS. La señal original presenta fluctuaciones en la línea de base debido a ruido de baja frecuencia. Este ruido es evidente en los cambios irregulares de la línea que afectan la estabilidad de la señal. Al aplicar el método se observa que se ha eliminado el ruido de baja frecuencia y además los picos de la señal que corresponden a la actividad cardíaca se mantienen claramente visibles y no se ven alterados por el proceso de suavizado.

En la figura 1 las variaciones entre 4 y 8 segundos, se ve que la señal suavizada en azul sigue de cerca los cambios importantes de la señal original, pero sin las oscilaciones de baja frecuencia. Esto confirma que el Robust LOESS está ajustando la señal localmente sin afectar los datos importantes en esas secciones.

B. Non-Local Means (NLM)

Este método se utiliza para eliminar ruidos de alta frecuencia residuales después de haber aplicado otros filtros anteriormente, funciona encontrando y promediando patrones similares dentro de la señal. Primero selecciona una ventana de referencia para cada dato después busca, en toda la señal, otras ventanas que tengan un patrón similar a la ventana de referencia. Una vez que se encuentran ventanas similares, el algoritmo asigna un peso a cada una de estas ventanas en función de su similitud con la ventana de referencia. Cuanto más parecida sea una ventana al patrón de referencia, mayor será su peso. Luego el algoritmo toma un promedio, dándole más importancia a los puntos que provienen de ventanas con patrones similares y menos importancia a los que son diferentes. Al promediar ventanas similares de distintas partes de la señal, el NLM logra reducir el ruido sin eliminar los detalles importantes.[2]

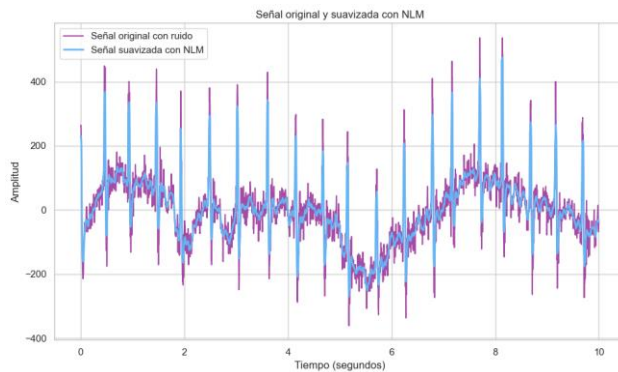


Figura 2: Señal ECG aplicando el método Non-Local Means

Se toma la misma señal ECG la cual presenta fluctuaciones fuertes e irregulares, características típicas de ruido. Este ruido afecta tanto la amplitud como la frecuencia de los picos de la señal. Al aplicar el método Non-local Means se observa que atenúa las variaciones abruptas. A pesar de la reducción de ruido, el método NLM ha conservado la forma general de la señal, incluyendo las oscilaciones naturales y los picos importantes.

Por otra parte, en la figura 3 se aplica tanto el método Robust LOESS como el método Non-Local Means, eliminando la distorsión de la línea base y ruido de alta frecuencia respectivamente. Los filtros se complementan logrando así una señal más limpia y cuidando los detalles del ECG.

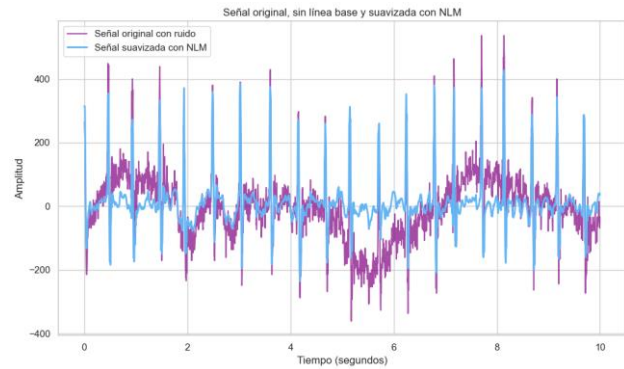


Figura 3: Señal ECG aplicando Robust LOESS y NLM.

III. FLUJOS DE PROCESAMIENTO DE DATOS

A. Filtro Wavelet

Investigaciones en el área del procesamiento de señales electrocardiográficas han demostrado que la familia de wavelets Daubechies ofrece un excelente desempeño para el análisis de señales ECG, debido a su capacidad para realizar una descomposición eficiente, segmentación precisa y reconstrucción fiel de la señal. Estas características son especialmente útiles en la identificación de características específicas como los picos R, la duración del complejo QRS, o las variaciones de la frecuencia cardíaca, ya que las wavelets Daubechies proporcionan una buena resolución tanto en el dominio del tiempo como en el de la frecuencia. Además, su adaptabilidad a diferentes escalas permite una representación multiresolución, lo que facilita la detección de patrones sutiles en las señales, tales como arritmias o anomalías en el ritmo cardíaco. Esta flexibilidad y precisión en la representación de señales ECG complejas la hace especialmente valiosa en la mejora de diagnósticos automáticos y la reducción de artefactos en el análisis clínico [3].

En varios estudios se ha recomendado el uso de la familia wavelet 'db4' con un rango de niveles de descomposición entre 2 y 4, debido a su capacidad para descomponer eficientemente las señales ECG y mantener las características relevantes de la señal. Sin embargo, otros artículos sugieren que es posible emplear hasta 6 niveles de descomposición. Adicionalmente, se utiliza un enfoque de umbralización 'soft', que prioriza la reducción de ruido mientras mantiene la integridad de las características principales de la señal. Este tipo de umbralización permite modificar los coeficientes de la señal, eliminando componentes de alta frecuencia y artefactos sin afectar demasiado las componentes importantes de la señal ECG, como los complejos QRS o los picos R[4]. Por otra parte, un estudio específico menciona que la combinación de 'db4' con nivel 4 ha demostrado ser la más efectiva, mostrando la mejor capacidad para eliminar el ruido y proporcionando una respuesta optimizada en términos de

preservación de la información relevante de la señal ECG. Estos resultados subrayan la importancia de elegir tanto la familia wavelet como el número de niveles adecuados, además de la estrategia de umbralización, para lograr un filtrado eficaz y una señal reconstruida de alta calidad [8].

En este proyecto, se implementa un filtro wavelet utilizando la familia 'db4' y un nivel de descomposición de nivel 4. Primero, se calcula la Transformada Wavelet Discreta (DWT) de la señal, que descompone la señal en coeficientes de aproximación y detalle. Luego, se aplica un umbral basado en la mediana de los coeficientes de detalle más finos, lo que se conoce como umbralización 'soft'. Este tipo de umbral suaviza los coeficientes al reducir aquellos menores que el umbral a cero y atenuar los más grandes, con el fin de eliminar el ruido presente en la señal. Después, los coeficientes modificados son reconstruidos usando la Transformada Wavelet Discreta Inversa (IDWT), obteniendo una nueva señal limpia de variaciones no deseadas y tendencias.

B. Flujo 1

El primer paso en el procesamiento de la señal es la aplicación de un filtro pasa-altas utilizando un filtro IIR (Infinite Impulse Response) a una frecuencia de corte de 0.5 Hz. Este filtro es esencial para eliminar las componentes de baja frecuencia que pueden ser causadas por el ruido de base, como los artefactos respiratorios o movimientos del paciente. La elección de 0.5 Hz como frecuencia de corte se basa en el comportamiento típico de las señales de ECG, donde las frecuencias inferiores a este valor no aportan información relevante y pueden interferir con la calidad de la señal. El uso de un filtro IIR, específicamente un filtro de Chebyshev tipo II, se justifica por su capacidad para lograr una atenuación efectiva en la banda de rechazo sin introducir distorsiones notables en la señal [5]. La orden del filtro, 5, asegura una transición suave entre las bandas pasante y de rechazo según lo reportado en la literatura donde se demuestra que con este orden del filtro es capaz de funcionar correctamente [6].

El segundo paso implica la aplicación de un filtro basado en la transformada wavelet discreta, explicada anteriormente.

Finalmente, se aplica un filtro pasa-bajas de 50 Hz para eliminar el ruido de alta frecuencia, como el artefacto electromagnético generado por equipos electrónicos cercanos. En este caso, se utiliza un filtro FIR (Finite Impulse Response) con una frecuencia de corte de 50 Hz. El uso de un filtro FIR se justifica por su estabilidad y la capacidad de obtener una respuesta de frecuencia lineal en la banda pasante, lo cual es esencial para mantener la fidelidad de la señal. Además, se opta por un filtro FIR con una ventana Kaiser, que permite un control preciso sobre la transición entre las bandas pasante y de rechazo mediante el parámetro beta [5]. Se seleccionó una orden de 250 para garantizar una atenuación adecuada fuera de la banda pasante sin

comprometer la precisión temporal de la señal dado que en la literatura se encontró que en este orden se lograba eliminar la contaminación de ruido [7].

C. Flujo 2

El primer paso en el flujo de procesamiento es el detrending, que consiste en remover una tendencia lineal de la señal. Las señales biomédicas, como las de ECG, a menudo presentan una deriva o desplazamiento de bajo nivel debido a cambios lentos en las condiciones del dispositivo o el paciente. Esta tendencia puede dificultar el análisis de las variaciones rápidas de la señal, que son las que contienen la información más relevante. La función detrend aplicada a la señal elimina esta tendencia, asegurando que se concentren los análisis en las fluctuaciones de interés.

El segundo paso implica la aplicación de un filtro basado en la transformada wavelet discreta, explicada anteriormente.

Finalmente, se aplica un filtro pasa-bajas con una frecuencia de corte de 50 Hz. Este paso es necesario para eliminar el ruido de alta frecuencia, especialmente el artefacto de la red eléctrica, que se presenta como interferencia a frecuencias cercanas a 50 Hz o 60 Hz.

Para este caso, se utiliza un filtro IIR (Infinite Impulse Response) con el filtro Chebyshev tipo II [5]. La elección de un filtro IIR se debe a su eficiencia en términos de número de coeficientes necesarios para lograr una atenuación adecuada en la banda de rechazo. Además, el filtro IIR tiene una respuesta más eficiente en términos de recursos computacionales para alcanzar la atenuación deseada. La frecuencia de corte de 50 Hz es estándar para este tipo de señales, ya que cubre el rango de interferencia electromagnética típicamente presente en las señales biomédicas.

El filtro de tipo Chebyshev tipo II se selecciona debido a su alta atenuación en la banda de rechazo, lo cual es crucial para suprimir los artefactos de alta frecuencia. La orden del filtro se ha establecido en 5 para equilibrar entre complejidad y desempeño, asegurando que el filtro no introduzca distorsiones significativas en la señal procesada [6].

D. Flujo 3

El primer paso en el procesamiento de la señal es la aplicación de un filtro pasa-altas utilizando un filtro IIR con el propósito de usar el tipo Chebyshev Tipo II con una frecuencia de corte de 0.5 Hz.

Posteriormente, se utiliza un filtro FIR (Finite Impulse Response) con una frecuencia de corte de 50 Hz. El uso de un filtro FIR se justifica por su estabilidad y la capacidad de obtener una respuesta de frecuencia lineal en la banda pasante, lo cual es esencial para mantener la fidelidad de la señal. Además, se opta por un filtro FIR con una ventana Kaiser [5].

E. Análisis y selección del flujo de procesamiento

Se seleccionaron al azar un total de 10 sujetos, divididos en dos grupos de 5 sujetos: el primer grupo con diagnóstico de Bradicardia Sinusal (SB) y el segundo con Fibrilación Auricular (AFIB). A cada uno de estos sujetos se le aplicaron los tres flujos de procesamiento propuestos, generando un conjunto de datos representado en la Figura 4 para los sujetos con SB y la Figura 5 para los sujetos con AFIB.

El objetivo de este procedimiento es comparar y analizar el desempeño de cada flujo de procesamiento en función de su capacidad para atenuar el ruido y preservar las características fisiológicas específicas de cada tipo de señal (SB y AFIB). Esto permitirá identificar el flujo de procesamiento que proporciona la mayor claridad y fidelidad en la señal para cada condición.

En la figura 4, se muestra la señal ECG de cinco sujetos con SB procesada mediante tres flujos de procesamiento, se pueden notar varias características distintivas entre los flujos en términos de claridad de señal y reducción de artefactos. En términos de comparación, el Flujo 2 ofrece la señal con la línea base más estable, pero a costa de cierta atenuación en la morfología de la señal. El Flujo 3 es el que parece preservar mejor la forma de la señal ECG, aunque con ligeras oscilaciones residuales. El Flujo 1, mientras tanto, proporciona una atenuación del ruido razonable y buena visibilidad de los picos, pero con una línea base menos estable que en el Flujo 2.

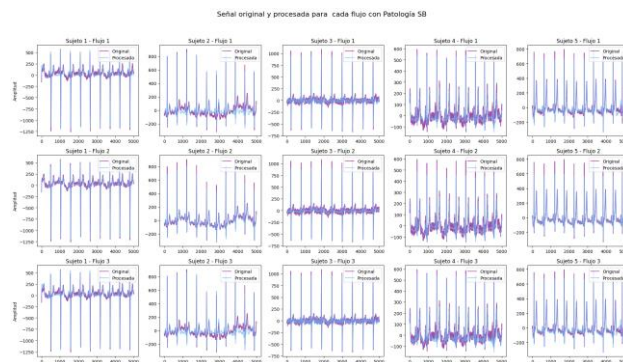


Figura 4: Señales ECG de 5 sujetos con SB para cada uno de los tres flujos de procesamiento.

En la figura 5, se muestra la señal ECG de cinco sujetos con AFIB procesada mediante tres flujos de procesamiento el Flujo 2 es el más efectivo para estabilizar la línea de base, lo cual es útil en AFIB donde la deriva podría interferir en el análisis. El Flujo 3 es el que mejor preserva la morfología de la señal y la variabilidad entre picos, siendo ideal para análisis morfológicos detallados. El Flujo 1, en cambio, ofrece una opción balanceada entre reducción de ruido y preservación de la señal, aunque con menos estabilidad en la línea de base.

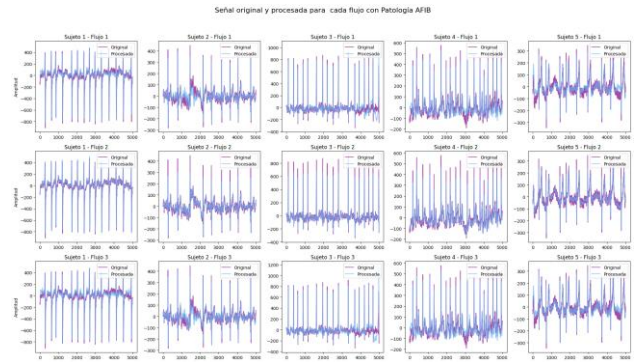


Figura 5: Señales ECG de 5 sujetos con AFIB para cada uno de los tres flujos de procesamiento.

Se seleccionó un sujeto representativo de cada grupo: uno con SB y otro con AFIB. A ambos se les aplicaron los tres flujos de procesamiento, permitiendo así visualizar directamente las diferencias en la calidad de la señal procesada y la efectividad de cada método.

Los resultados obtenidos en la figura 6 y 7 muestran de forma clara que el Flujo 2 es el más efectivo en la estabilización de la línea de base y la reducción de artefactos de baja frecuencia. Esto es particularmente notable en las señales de AFIB, donde las variaciones irregulares de los intervalos entre picos pueden verse comprometidas por el ruido de baja frecuencia y la deriva de la línea de base. En comparación con los otros flujos, el Flujo 2 logra una señal más limpia y uniforme, sin sacrificar la claridad de los picos característicos de la señal ECG, tanto en condiciones de SB como de AFIB. Esta estabilidad es crucial en contextos de análisis de la morfología de la señal, especialmente en patologías como la AFIB, donde es necesario identificar variaciones sutiles entre latidos.

El análisis reafirma las observaciones previas, destacando que el Flujo 2 ofrece el mejor equilibrio entre la atenuación de artefactos y la preservación de la morfología de la señal.



Figura 6: Señal ECG del sujeto 1 con SB para cada uno de los flujos de procesamiento.

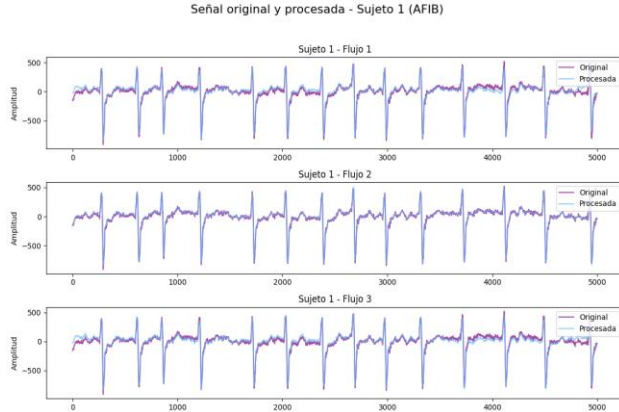


Figura 7: Señal ECG del sujeto 1 con AFIB para cada uno de los flujos de procesamiento.

Aunque el análisis gráfico permite una comparación visual preliminar de la efectividad de cada flujo de procesamiento de señal, este enfoque es subjetivo y limitado para establecer conclusiones sólidas. La variabilidad en la interpretación visual y las limitaciones inherentes a la observación subjetiva hacen necesario un análisis cuantitativo que permita identificar de forma objetiva el flujo de procesamiento más adecuado. Para abordar esta necesidad, se implementó una evaluación cuantitativa basada en el error cuadrático medio (MSE) entre la señal original y la señal procesada para cada flujo.

El algoritmo de selección de flujo utiliza el MSE como métrica de error para cuantificar la diferencia entre la señal original y la señal procesada. A cada uno de los cinco sujetos en los grupos SB (Sinusal Bradirritmia) y AFIB (Fibrilación Auricular) se les aplicaron los tres flujos de procesamiento. Para cada sujeto, se calcularon los errores individuales en los tres flujos (Flujo 1, Flujo 2 y Flujo 3), y estos errores fueron almacenados en un diccionario para facilitar su análisis.

Una vez calculados los errores para cada sujeto en los tres flujos, se obtuvieron los errores promedio de cada flujo (es decir, el promedio de los cinco valores de MSE para cada flujo). Esta medida de error promedio representa la efectividad general de cada flujo en la preservación de la forma de la señal original tras el procesamiento.

Flujo	Error Promedio
1	1290.64
2	571.53
3	1245.98

Tabla 1: Resultados del error promedio para cada flujo de procesamiento.

Finalmente, el flujo con el menor error promedio fue seleccionado como el flujo ideal, ya que un menor MSE

indica una mayor similitud entre la señal original y la señal procesada, lo cual sugiere que dicho flujo reduce efectivamente el ruido sin introducir distorsiones significativas en la morfología de la señal ECG. Los resultados de este análisis cuantitativo muestran que el Flujo 2 presenta el menor error promedio en comparación con los otros dos flujos, confirmando así que es el más adecuado para preservar la estructura de la señal ECG tanto en sujetos con SB como con AFIB.

IV. NORMALIZACIÓN DE DATOS

El artículo describe que la mayoría de los conjuntos de datos de ECG muestran tendencias generales que podrían surgir de los movimientos musculares del paciente o de las fluctuaciones en el voltaje del aparato de registro. En consecuencia, resulta imperativo erradicar estas tendencias antes de avanzar con los procedimientos de procesamiento de señales. Para lograr esto, se emplea una metodología de ajuste polinomial denominada Detrend.

Para facilitar una comparación precisa de las señales sin sesgos, es imperativo que todas las señales estén dentro de un rango estandarizado antes del examen. Este procedimiento de normalización se lleva a cabo mediante una técnica denominada *compresión* descrita por la ecuación 1, que implica la modificación de los valores de todas las señales para ajustarlas a un rango uniforme, preservando así sus atributos esenciales y mitigando las fluctuaciones externas de magnitud. Esta compresión permite realizar una evaluación equitativa entre los distintos conjuntos de datos, garantizando que las diferencias detectadas se deriven de variaciones intrínsecas dentro de la señal y no de discrepancias en la escala de los datos.

$$s(t_k) = \frac{c(t_k) - x_{min}}{x_{max} - x_{min}} \quad [1]$$

El proceso de eliminación de tendencias Detrend se repite en este momento para eliminar cualquier tendencia residual que pueda persistir en las señales del electrocardiograma (ECG) tras las transformaciones anteriores. Esta segunda implementación de la metodología de ajuste polinomial es imprescindible para garantizar que los atributos de las señales permanezcan inalterados por variaciones globales externas.

V. FRECUENCIA MÁXIMA

Se calcula el espectro de potencia de una señal, y se extrae la frecuencia que contiene la mayor potencia usando el método de Welch.

Se establece un valor máximo de nperseg de 512. Esto asegura que las ventanas sean suficientemente largas para capturar la dinámica de la señal sin perder resolución en

frecuencia. Este tamaño de ventana es común en análisis de señales de ECG y resulta en una buena resolución temporal-frecuencia para las bandas de interés típicas de esta señal.

Se establece un solapamiento del 75%, este solapamiento es común en el método de Welch, ya que permite suavizar la variabilidad de la potencia entre ventanas consecutivas y mejorar la resolución en frecuencia del espectro estimado sin perder información crítica de la señal. Un solapamiento elevado mejora la estabilidad del espectro de potencia calculado. Adicionalmente, se emplea una ventana de Hanning para reducir los efectos de los extremos de cada segmento de ventana, minimizando así el "leakage" espectral o fuga de energía.

En el proyecto, además del método de Welch, también se utiliza la Transformada Rápida de Fourier (FFT) para calcular el espectro de potencia de la señal. En el Proyecto 2, el cálculo del espectro de potencia se realizó exclusivamente mediante la FFT, mientras que en el presente proyecto (Proyecto 3) se incluyen tanto la FFT como el método de Welch. Esta adición permite comparar y validar los resultados de ambos métodos de análisis espectral, y optimizar la precisión en la identificación de frecuencias de interés en las señales de ECG. La comparación entre ambos enfoques facilita una evaluación más robusta de la potencia de frecuencia en los datos y permite concluir si el flujo de procesamiento en el Proyecto 3 ofrece ventajas sobre el Proyecto 2 en términos de precisión y resolución espectral.

El método de Welch y la Transformada Rápida de Fourier (FFT) difieren principalmente en la forma en que procesan y presentan el espectro de potencia de una señal. Welch divide la señal en segmentos solapados y promedia sus espectros, lo que reduce la varianza y suaviza el espectro, haciéndolo más estable y menos sensible al ruido. Esto resulta útil para analizar señales no estacionarias o ruidosas, como el ECG. Por otro lado, la FFT calcula el espectro en una sola pasada sin segmentación, lo que permite una resolución en frecuencia más alta y precisión en la identificación de frecuencias exactas, pero es más sensible al ruido y puede ser menos representativa en señales variables. Aunque la FFT es computacionalmente más rápida, Welch ofrece una mayor robustez en la interpretación de señales complejas.

Al aplicar la rutina se obtiene el siguiente dataframe, donde se evidencia la $fMP(FFT)$ y $fMP(Welch)$.

Se observan algunas discrepancias en estos cálculos que pueden deberse principalmente a las diferencias en cómo cada método maneja la resolución espectral y el ruido. La FFT tiene una alta resolución espectral y es más sensible a picos de alta frecuencia o transitorios en la señal, lo que puede generar diferencias notables, especialmente en señales con

componentes de frecuencia rápida, como en la fibrilación auricular. Por otro lado, Welch suaviza la señal al promediar los espectros de segmentos superpuestos, lo que reduce la influencia de los picos transitorios y el ruido, pero a costa de una menor resolución. Así, las discrepancias reflejan cómo cada método responde a las características y fluctuaciones de la señal.

	Archivo	Estado	fMP(FFT)	fMP(welch)
0	MUSE_20180111_155154_74000.csv	AFIB	9.7	4.88281
1	MUSE_20180111_155203_15000.csv	AFIB	4.2	3.90625
2	MUSE_20180111_155758_17000.csv	SB	0.9	3.90625
3	MUSE_20180111_155839_67000.csv	SB	1.0	0.97656
4	MUSE_20180111_155859_36000.csv	SB	0.9	0.97656
...
5664	MUSE_20180712_161449_58000.csv	AFIB	0.2	0.97656
5665	MUSE_20180712_161451_72000.csv	AFIB	0.1	0.97656
5666	MUSE_20180712_161452_78000.csv	AFIB	0.3	0.97656
5667	MUSE_20180712_161503_00000.csv	AFIB	0.2	3.90625
5668	MUSE_20180712_161505_16000.csv	AFIB	0.1	0.97656

Figura 8: Dataframe Final

VI. ANÁLISIS PROYECTO 3

A. Estadística

Se realiza un diagrama de caja y bigotes donde se compara la frecuencia de máxima potencia (fMP) entre fibrilación auricular (AFIB) y bradicardia sinusal (SB) como el de la figura 9.

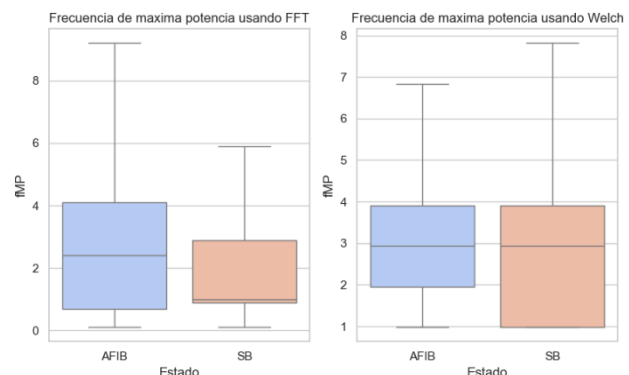


Figura 9: comparación entre FFT y Welch

En la figura 9 usando FFT se observa que para AFIB la mediana de la frecuencia de máxima potencia se sitúa alrededor de 2, con un rango intercuartílico (IQR) que va aproximadamente de 1 a 4. Para SB La mediana es un poco menor que la de AFIB, situándose cerca de 1,5. El IQR va aproximadamente de 1 a 3, estos resultados muestran que el estado AFIB tiene una distribución más amplia en comparación con SB, lo que sugiere que en el estado AFIB puede haber una mayor variabilidad en la frecuencia de máxima potencia cuando se usa la FFT.

Por otra parte, utilizando Welch se observa que en AFIB La mediana de fMP se sitúa un poco más alta en comparación con la FFT, alrededor de 3, con un IQR que va aproximadamente de 2 a 4. Para SB la mediana es similar a

la del estado AFIB, también alrededor de 3, y el IQR es similar, entre 2 y 4. A diferencia del método FFT, aquí los estados AFIB y SB muestran una distribución muy similar en cuanto a la mediana y la dispersión de la frecuencia de máxima potencia.

También se realizan pruebas de normalidad y homocedasticidad para las frecuencias de máxima potencia (fMP) obtenidas mediante los métodos FFT y Welch en dos estados, SB y AFIB. Los valores p de normalidad en ambos métodos son extremadamente bajos, lo que indica que las distribuciones de las frecuencias de máxima potencia no siguen una distribución normal en ninguno de los casos.

En cuanto a la homocedasticidad, que evalúa si las varianzas de los datos son iguales entre los grupos SB y AFIB, los resultados muestran una clara diferencia entre los métodos. Para FFT, el valor p es $5.85e-66$, lo que indica heterocedasticidad; es decir, las varianzas entre los grupos no son iguales. Por otro lado, el método Welch presenta un valor p de 0.2107, que es superior al umbral de 0.05, lo que sugiere homocedasticidad y, por tanto, varianzas similares entre los estados SB y AFIB.

Por tanto, se concluye que el método Welch es más adecuado para el análisis de la frecuencia de máxima potencia en señales ECG, ya que proporciona una representación más estable y menos afectada por diferencias de varianza entre los estados (ruido). Esto hace que Welch sea más confiable para el estudio comparativo entre los estados AFIB y SB

	Metric	Mann-Whitney U (stat)	p-value	Significance
0	fMP(FFT)	2787542.5	2.971324e-32	Hay diferencia Significativa
1	fMP(Welch)	2785237.5	3.569328e-34	Hay diferencia Significativa

Figura 10: prueba de Mann-Whitney U para fMP

Dado que la distribución de los grupos no es normal se realiza la prueba de U-Mann-Whitney obteniendo que en los dos casos hay una diferencia significativa en la frecuencia de máxima potencia. Sin embargo, dado que la prueba de Mann-Whitney es una prueba no paramétrica y se confirma que no hay normalidad en los datos, este resultado respalda la robustez del método Welch para el análisis de ECG, debido a su estabilidad adicional observada en pruebas previas de homocedasticidad.

B. Estadística descriptiva

Se calcula la media, mediana y desviación estándar de las señales ECG para determinar si hay diferencia entre los sujetos con AFIB y SB.

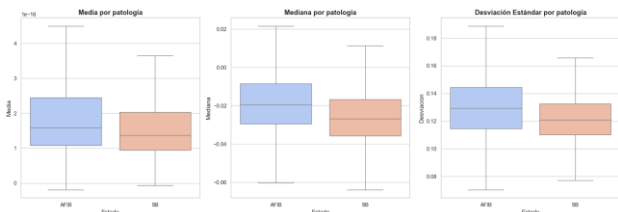


Figura 11: Media, mediana y desviación estándar entre los sujetos con AFIB y SB

En la figura 11 se observa que la media en el grupo AFIB presenta una mayor dispersión en comparación con SB. Además, el rango intercuartílico de la media es más amplio en AFIB que en SB, lo que sugiere una mayor variabilidad en los sujetos con fibrilación auricular. En cuanto al gráfico de la mediana, las medianas de ambos grupos son similares en magnitud; sin embargo, el grupo AFIB muestra una mayor dispersión de valores en comparación con SB, indicando una mayor variabilidad en la mediana en los sujetos con fibrilación auricular. Por último, en el gráfico de la desviación estándar, AFIB presenta un valor más alto que SB, lo cual reafirma que la variabilidad de la señal es mayor en los casos de fibrilación auricular.

Se realizan pruebas de normalidad y homocedasticidad para las métricas de media, mediana y desviación estándar en los grupos SB y AFIB. En el cual los resultados de valor p de normalidad para ambos grupos y para las tres métricas son bajos ($p < 0.05$), lo cual indica que los datos no siguen una distribución normal en ninguna de las métricas evaluadas para ambos grupos, AFIB y SB. Con respecto a la homocedasticidad los valores p también son bajos para las tres métricas (media, mediana y desviación estándar), indicando heterocedasticidad. Esto significa que las varianzas entre los dos grupos no son homogéneas, lo cual podría influir en la interpretación de diferencias entre los grupos.

Como las tres métricas distribuyen de forma no normal se realiza la prueba paramétrica de Mann-Whitney U para evaluar si existen diferencias significativas entre dos grupos, puesto que el valor p para las tres métricas es bajo (menores que 0.05), confirma una diferencia significativa entre SB y AFIB en esta métrica.

	Metric	Mann-Whitney U Statistic	Mann-Whitney p-value	Significancia
0	Media	2962364.0	2.723390e-18	Hay Diferencia Significativa
1	Mediana	2511302.0	5.968149e-62	Hay Diferencia Significativa
2	Desviacion	2728425.0	1.389017e-37	Hay Diferencia Significativa

Figura 12: prueba de Mann-Whitney U para media, mediana y desviación estándar

Hipótesis para cada métrica

- Media
- Hipótesis Nula (H_0): La media en el grupo SB es igual a la media en el grupo AFIB. No existe una

diferencia significativa entre las medias de ambos grupos.

- Hipótesis Alternativa (H_1): La media en el grupo SB es diferente a la media en el grupo AFIB. Existe una diferencia significativa entre las medias de ambos grupos.
- Mediana
 - Hipótesis Nula (H_0): La mediana en el grupo SB es igual a la mediana en el grupo AFIB. No existe una diferencia significativa entre las medianas de ambos grupos.
 - Hipótesis Alternativa (H_1): La mediana en el grupo SB es diferente a la mediana en el grupo AFIB. Existe una diferencia significativa entre las medianas de ambos grupos.
- Desviación Estándar
 - Hipótesis Nula (H_0): La desviación estándar en el grupo SB es igual a la desviación estándar en el grupo AFIB. No existe una diferencia significativa en la variabilidad entre los dos grupos.
 - Hipótesis Alternativa (H_1): La desviación estándar en el grupo SB es diferente a la desviación estándar en el grupo AFIB. Existe una diferencia significativa en la variabilidad entre los dos grupos.

Dado que los p-valores para las pruebas de Mann-Whitney en las tres métricas (media, mediana y desviación estándar) son extremadamente bajos (menores que 0.05), se rechaza la hipótesis nula para cada métrica y se acepta la hipótesis alternativa.

VII. ANÁLISIS PROYECTO 2 Y 3

Se realiza diagrama de caja y bigotes en el cual se compara la frecuencia de máxima potencia (fMP) entre los proyectos dos y tres para los estados de AFIB y SB. Utilizando dos métodos distintos FFT y Welch

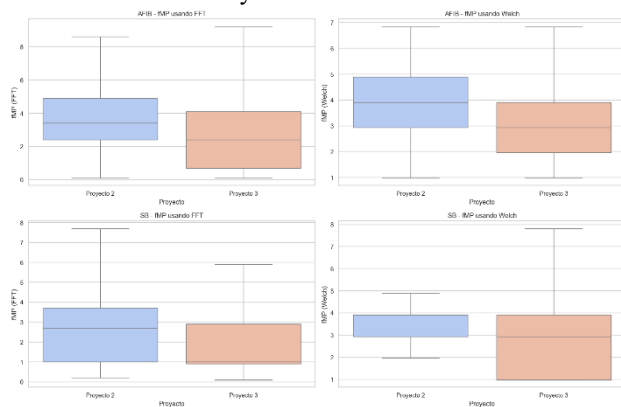


Figura 13: Comparación de fMP para el proyecto 2 y proyecto 3

En la figura 13 al analizar para AFIB con respecto al proyecto dos presenta una mayor dispersión de valores de fMP, con una mediana cercana a 3.5, en contraste con el proyecto tres presenta una mediana ligeramente menor, alrededor de 3, esto sugiere que el proyecto dos en AFIB con FFT tiene una mayor variabilidad en fMP, mientras que el proyecto tres muestra valores más concentrados alrededor de su mediana. Para AFIB utilizando Welch ambos proyectos tiene valores medianos relativamente altos.

Ahora para SB usando FFT el proyecto dos presenta una mediana de fMP en torno a 3, con una dispersión moderada por otra parte el proyecto tres en esta métrica tiene una mediana ligeramente menor, cerca de 2.5 y presenta menor dispersión. En este caso, al igual que en AFIB, Proyecto 2 muestra una mayor variabilidad en los valores de fMP en comparación con Proyecto 3. ahora Sb usando Welch el proyecto 3 presenta más variabilidad en los valores de fMP en comparación con el proyecto 2, al contrario de lo que se observa en FFT.

El análisis sugiere que Proyecto 2 presenta una mayor variabilidad en fMP, especialmente en el estado AFIB y con el método Welch. El proyecto 3, por otro lado, muestra valores de fMP más concentrados y estables, especialmente en el estado SB y usando el método FFT. Esto sugiere que las características del Proyecto 3 podrían estar asociadas con una menor variabilidad en frecuencia de máxima potencia.

Estado	Metric	Mann-Whitney U Statistic	Mann-Whitney p-value	Significancia
0	AFIB	Media	1301737.5	3.206797e-20 Hay Diferencia Significativa
1	SB	Media	5771070.0	3.909018e-73 Hay Diferencia Significativa
2	AFIB	Mediana	1297616.0	9.083952e-21 Hay Diferencia Significativa
3	SB	Mediana	5443041.0	1.291151e-101 Hay Diferencia Significativa
4	AFIB	Desviacion	1307523.0	1.828011e-19 Hay Diferencia Significativa
5	SB	Desviacion	5480467.0	3.988340e-98 Hay Diferencia Significativa

Figura 14: Prueba U-Mann Whitney para la estadística descriptiva entre el proyecto 2 y proyecto 3.

Al no seguir una distribución normal se realiza la de Mann-Whitney U para comparar las métricas de media, mediana y desviación estándar en los estados AFIB y SB. Para poder determinar diferencias significativas entre los dos grupos.

Los resultados de la prueba de Mann-Whitney indican que existen diferencias estadísticamente significativas entre los proyectos para cada métrica (mediana, media y desviación estándar) en ambos estados (AFIB y SB). Esto sugiere que los proyectos difieren en cuanto a sus valores promedio, su valor central y su dispersión en cada estado de salud evaluado. La presencia de diferencias significativas en estas tres métricas respalda la conclusión de que los datos de los proyectos en cada estado no son homogéneos y, por tanto, presentan características distintas en los grupos de sujetos con AFIB y SB.

Para corroborar esto también se calculó los valores cuadrático medio (ECM) entre la frecuencia de máxima potencia (fMP)

calculada con dos métodos diferentes, Welch y FFT, para comparar los datos entre el Proyecto 2 y el Proyecto 3. Como resultado el método Welch muestra un ECM más bajo en comparación con FFT al analizar la fMP entre los proyectos 2 y 3, lo cual indica que los valores de fMP obtenidos mediante Welch son más similares entre los dos proyectos. Esto sugiere que el método Welch proporciona resultados más consistentes y menos sensibles a las diferencias entre proyectos en comparación con FFT, lo que podría hacer de Welch una opción más adecuada para el análisis de la fMP.

Por otro lado, al realizar un análisis visual comparativo entre la señal prefiltrada utilizada en el proyecto 2 y la señal procesada en el proyecto 3, para un sujeto con SB y otro con AFIB, se observa que el flujo de procesamiento 2 se ajusta de manera bastante precisa a las características de la señal prefiltrada. Sin embargo, aunque la señal procesada muestra una mejora considerable, todavía presenta algunos niveles de ruido y contaminación que impiden obtener una señal completamente limpia. Esto sugiere que el flujo de procesamiento actual, si bien efectivo, no es suficiente para garantizar una eliminación total de los artefactos presentes en las señales ECG. Por lo tanto, para futuras investigaciones, se recomienda explorar un flujo de procesamiento más agresivo que incorpore técnicas adicionales de reducción de ruido, como Robust LOESS (Locally Weighted Scatterplot Smoothing) y Non-Local Means (NLM). Estas metodologías, complementadas con el flujo ya establecido, podrían ofrecer una mayor efectividad en la atenuación de ruido, mejorando así la calidad de las señales procesadas y permitiendo un análisis más preciso y fiable en el diagnóstico de arritmias cardiovasculares.

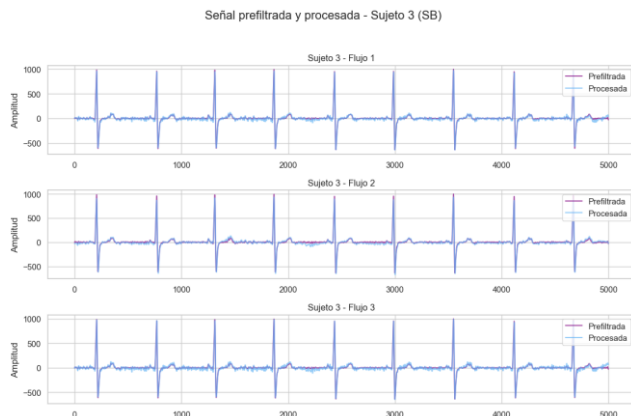


Figura 15: Comparación de la señal prefiltrada (proyecto2) y la señal procesada (proyecto 3) para un sujeto con SB.

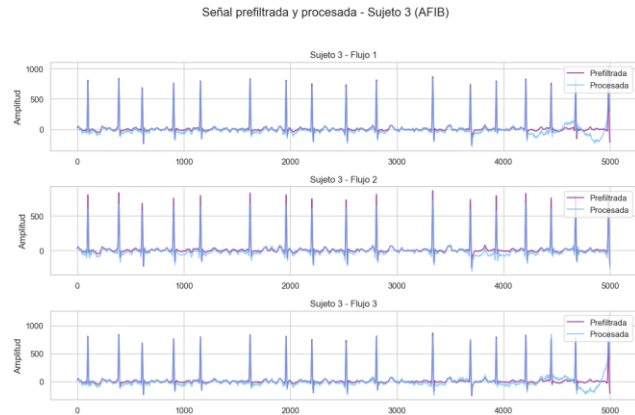


Figura 16: Comparación de la señal prefiltrada (proyecto2) y la señal procesada (proyecto 3) para un sujeto con AFIB.

VIII. CONCLUSIONES

- El flujo seleccionado de procesamiento es el número 2 que incluye detrend, filtro wavelet modificado y filtrado pasabajas a 50 Hz, mejora significativamente la calidad de las señales ECG. El detrend elimina las tendencias de baja frecuencia, mientras que el filtro wavelet modificado reduce el ruido de alta frecuencia sin perder detalles importantes. El filtrado pasabajas a 50 Hz elimina la interferencia de la red eléctrica, utilizando un filtro IIR. Este enfoque integrado asegura señales limpias y adecuadas para un análisis preciso y fiable en aplicaciones clínicas.
- Aunque el flujo de procesamiento 2 demuestra eficacia en la eliminación de la contaminación de la señal original, no logra una coincidencia total con las señales prefiltradas. Esta discrepancia podría estar relacionada con el uso de un filtro IIR pasabajas, el cual, aunque eficiente, puede introducir distorsiones de fase que afectan la calidad de la señal. Además, las características específicas del filtro wavelet utilizado pueden influir en la capacidad de aproximarse a las señales prefiltradas. Es posible también que las señales prefiltradas hayan seguido un proceso de filtrado distinto, lo que contribuye a que no se logre una igualdad del 100% entre ambas.
- Se observan diferencias significativas entre los resultados de fMP entre el proyecto 2 y el proyecto 3. Esto se debe a que la señal procesada en el proyecto 3 no se ajusta completamente a la señal prefiltrada utilizada en el proyecto 2, lo que sugiere que el flujo de procesamiento en el proyecto 3 aún no ha logrado una limpieza óptima de la señal.

- El análisis del error cuadrático medio (ECM) entre los resultados de fMP obtenidos con el método Welch y FFT en los proyectos 2 y 3 revela que, en ambos casos, existe una discrepancia notable entre las señales procesadas. El ECM más bajo se presenta en fMP(Welch) con un valor de 6.05, mientras que el ECM de fMP(FFT) es aún mayor, con un valor de 8.13.

REFERENCIAS

- [1] Zheng, J., Chu, H., Struppa, D., Zhang, J., Yacoub, S. M., El-Askary, H., ... & Rakovski, C. (2020). Optimal multi-stage arrhythmia classification approach. *Scientific reports*, 10(1), 2898. <https://www.nature.com/articles/s41598-020-59821-7>
- [2] Manjón, J. V., Carbonell-Caballero, J., Lull, J. J., García-Martí, G., Martí-Bonmatí, L., & Robles, M. (2008). MRI denoising using non-local means. *Medical image analysis*, 12(4), 514-523. <https://www.sciencedirect.com/science/article/pii/S1361841508000248>
- [3] García, J., Sosa, C., & Petrino, R. (n.d.). Filtrado wavelet sobre FPGA en tiempo real para la Corrección de Tendencias en ECG. ResearchGate. https://www.academia.edu/36117743/Filtrado_Wavelet_sobre_FPGA_en_Tiempo_Real_para_Correcci%C3%B3n_de_Tendencias_en_ECG
- [4] Gonzalez, M., & Martinez, V. (2017). Filtrado y caracterización de señales electrocardiográficas sobre R. ResearchGate. https://www.researchgate.net/publication/266855385_Procesamiento_Wavelet_de_senales_electrocardiograficas_sobre_R
- [5] Monroy, N., Villamizar, J., Otero, M., & Altuve, M. (2017). Análisis del desempeño de filtro IIR y FIR para la reducción del desplazamiento de la línea base del ECG en Matlab. VI Congreso Venezolano de Bioingeniería 2017, 154-157. https://www.researchgate.net/publication/324416137_Analisis_del_desempeno_de_filtros_IIR_y_FIR_para_la_reduccion_del_desplazamiento_de_la_linea_de_base_del_ECG_en_Matlab
- [6] Tibamoso, G. (2007). Reducción de Interferencias en Señales ECG mediante Filtros Digitales IIR. *I+D*, 5(2), 33-40. https://revistas.uptc.edu.co/index.php/ingenieria_sogamoso/article/view/858/857
- [7] Domper, B. (2018). Filtrado de la señal de ritmo cardíaco basado en la transformada Hilbert. Universitat de Lleida.
- [8] Tovar, D. A., Orozco, A. J., & Muñoz, P. A. (n.d.). Transformada wavelet y máquinas de soporte vectorial para la identificación de arritmias. *rev.invest.univ.quindio*, 19, 104-114.