



UNIVERSIDAD DE LOS ANDES

FACULTAD DE CIENCIAS

DEPARTAMENTO DE FÍSICA

MÉRIDA – VENEZUELA

**ESTUDIO DE UN SISTEMA DE SEGUIMIENTO OCULAR
BASADO EN TÉCNICAS DE
PROCESAMIENTO DE IMÁGENES**

Una Propuesta Experimental

Trabajo Especial de Grado presentado
como requisito parcial para optar al título de

LICENCIADO EN FÍSICA

Autor:

Víctor Toro

Jurados:

Dr. John Ferreira

Profa. Stephanie Carrillo

Tutor Académico:

Prof. Juan Carlos Villegas

MÉRIDA, ENERO DE 2026

*A mis padres, por ser el pilar fundamental de mi vida y
por su apoyo incondicional en cada paso de este camino.*

A Dios, por la fortaleza espiritual.

*A mí mismo, por la perseverancia y la resiliencia
demostrada para alcanzar esta meta.*

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todas las personas e instituciones que hicieron posible la realización de este Trabajo Especial de Grado.

En primer lugar, a la Universidad de Los Andes y al Departamento de Física, mi casa de estudios, por haberme brindado la formación académica y las herramientas necesarias para mi desarrollo profesional.

A mi tutor, Prof. Juan Carlos Villegas, por su invaluable guía, paciencia y conocimientos compartidos durante todo el proceso de investigación. Gracias por confiar en esta propuesta y por sus acertadas correcciones que enriquecieron sustancialmente este trabajo.

A la Profa. Stephanie Carrillo, por su apoyo constante, sus consejos académicos y su disposición para colaborar en la revisión de los avances de esta tesis.

A mis padres y familiares, por su amor, sacrificio y comprensión durante mis años de estudio. Este logro es también suyo.

A mis compañeros y amigos de la carrera, con quienes compartí largas horas de estudio, retos y alegrías. Gracias por hacer el camino más llevadero.

Finalmente, a todos aquellos que de una u otra forma contribuyeron con un granito de arena para culminar esta etapa de mi vida.

Víctor Toro

Resumen

El ojo es una fuente de información valiosa que, al ser analizada rigurosamente, permite no solo obtener resultados biométricos únicos, sino también aplicar estos hallazgos en áreas diversas. Bajo esta premisa, esta investigación desarrolló un sistema de seguimiento ocular de bajo costo capaz de identificar personas y controlar interfaces mediante la mirada.

La metodología empleó un enfoque híbrido que combina características físicas con patrones dinámicos de movimiento (velocidad y precisión). Las pruebas con 15 participantes arrojaron una exactitud de identificación del 83.5 %, superando a los métodos que solo utilizan imágenes estáticas. Como aplicación práctica, se validó la viabilidad técnica del sistema para el control de cursor en tecnologías asistivas.

Adicionalmente, el análisis reveló que los patrones oculares varían según el nivel de cansancio e iluminación, sugiriendo un nuevo uso potencial para el monitoreo de fatiga en salud ocupacional. En conclusión, este trabajo demuestra que es factible implementar soluciones avanzadas de biometría y accesibilidad sin depender de hardware costoso.

Palabras clave: Seguimiento ocular, biometría, procesamiento de imágenes, tecnologías asistivas, inteligencia artificial.

Índice general

| | |
|---|------------|
| Agradecimientos | II |
| Resumen | III |
| Lista de Figuras | XII |
| Lista de Tablas | XIV |
| Introducción | XV |
| 1. Planteamiento del Problema | 1 |
| 1.1. El problema | 1 |
| 1.2. Objetivos | 2 |
| 1.2.1. Objetivo general | 2 |
| 1.2.2. Objetivos específicos | 2 |
| 1.3. Justificación | 2 |
| 1.4. Organización de la tesis | 3 |
| 2. Marco Teórico | 4 |
| 2.1. Antecedentes teóricos de la investigación | 4 |
| 2.1.1. Eye Aspect Ratio (EAR) para Detección de Somnolencia | 4 |
| 2.1.2. Introduction to Eye Tracking: A Hands-On Tutorial | 5 |
| 2.1.3. Remote Photoplethysmography | 5 |
| 2.1.4. AI in Medical Imaging Technology | 5 |
| 2.1.5. Abnormal Ocular Movement in Multiple-System Atrophy | 5 |
| 2.1.6. Aplicación de la convolución de matrices | 6 |
| 2.1.7. Analysis of eye-tracking experiments on Tobii T60 | 6 |
| 2.2. Bases conceptuales | 6 |
| 2.2.1. Anatomía del Ojo Humano | 6 |
| 2.2.2. Movimientos oculares básicos | 7 |
| 2.2.3. La secuencia principal (Main Sequence) | 8 |
| 2.3. Sistema de captura y procesamiento | 9 |
| 2.3.1. Tecnologías de seguimiento ocular | 9 |

| | |
|---|-----------|
| 2.3.2. Fundamentos de la captura por video | 9 |
| 2.3.3. Preprocesamiento de la señal | 10 |
| 2.4. Caracterización biométrica y dinámica de la señal | 11 |
| 2.4.1. Biometría oculomotora | 12 |
| 2.4.2. Pupilometría dinámica y cognitiva | 13 |
| 2.4.3. Análisis de la calidad del movimiento: Jerk y suavidad | 13 |
| 2.4.4. Análisis de señales no lineales y complejidad fractal | 14 |
| 2.5. Técnicas de análisis y clasificación | 14 |
| 2.5.1. Estadística descriptiva | 15 |
| 2.5.2. Análisis Discriminante Lineal (LDA) | 15 |
| 2.5.3. Máquinas de Vectores de Soporte (SVM) | 16 |
| 2.5.4. Random Forest (Bosque Aleatorio) | 17 |
| 3. Marco Metodológico | 20 |
| 3.1. Diseño y Tipo de Investigación | 20 |
| 3.1.1. Enfoque Cuantitativo | 20 |
| 3.1.2. Diseño Experimental | 21 |
| 3.2. Sistema Experimental y Herramientas | 22 |
| 3.2.1. Hardware e Instrumentación | 22 |
| 3.2.2. Entorno y Condiciones de Iluminación | 25 |
| 3.2.3. Herramientas de Software | 26 |
| 3.2.4. Disponibilidad del Código | 27 |
| 3.2.5. Arquitectura del Flujo de Datos (Pipeline) | 28 |
| 3.3. Protocolo de Adquisición de Datos | 30 |
| 3.3.1. Población y Muestra | 30 |
| 3.3.2. Diseño del Estímulo Visual | 33 |
| 3.3.3. Procedimiento Experimental | 34 |
| 3.4. Procesamiento Digital de Imágenes | 37 |
| 3.4.1. Preprocesamiento | 37 |
| 3.4.2. Segmentación y Detección de Pupila | 38 |
| 3.4.3. Estimación del Vector de Mirada (Gaze Vector) | 41 |
| 3.5. Procesamiento de Señales y Extracción de Métricas | 44 |
| 3.5.1. Cálculo de la Cinemática Angular | 44 |
| 3.5.2. Filtrado y Derivación | 44 |
| 3.5.3. Nuevas Métricas Biométricas Integradas | 45 |
| 3.5.4. Justificación y Selección del Espacio de Características | 46 |

| | |
|---|-----------|
| 3.6. Resumen del Capítulo | 49 |
| 4. Resultados y Discusión | 50 |
| 4.1. Validación del Sistema de Captura y Procesamiento | 50 |
| 4.1.1. Calidad de la Señal y Filtrado | 50 |
| 4.1.2. Precisión de la Detección (YOLOv8) | 54 |
| 4.2. Caracterización Cinemática y Fisiológica | 54 |
| 4.2.1. Análisis de la Secuencia Principal (Main Sequence) | 55 |
| 4.2.2. Perfiles de Velocidad y Jerk | 58 |
| 4.3. Identificación de Patrones Biométricos | 59 |
| 4.3.1. Importancia de Características (<i>Feature Importance</i>) | 60 |
| 4.3.2. Jerarquía de Relevancia Biométrica | 62 |
| 4.3.3. Perfiles Biométricos Individuales | 63 |
| 4.3.4. Visualización de Separabilidad (LDA) | 64 |
| 4.3.5. Evaluación de métodos alternativos: reducción dimensional y clustering | 65 |
| 4.4. Rendimiento de la Clasificación | 69 |
| 4.4.1. Métricas de los Modelos | 69 |
| 4.4.2. Análisis de Confusión y Escalabilidad del Clasificador | 70 |
| 4.5. Evaluación para Control de Cursor | 76 |
| 4.5.1. Mapeo y Corrección (Matriz de Homografía) | 76 |
| 4.5.2. Precisión Espacial y Estabilidad | 76 |
| 4.5.3. Viabilidad y Trabajo Futuro | 81 |
| 4.6. Análisis de Robustez Temporal y Deriva Biométrica | 81 |
| 4.6.1. Degradación del Rendimiento por Fatiga | 82 |
| 4.6.2. Análisis Estadístico de la Deriva: Ocho Métricas | 82 |
| 4.6.3. Mecanismo de Confusión: Convergencia Unidimensional | 83 |
| 4.6.4. Síntesis: Dominancia Unidimensional y Fragilidad del Sistema | 84 |
| 4.7. Discusión General | 85 |
| 5. Discusión, Conclusiones y Recomendaciones | 87 |
| 5.1. Discusión General de Resultados | 87 |
| 5.1.1. Cumplimiento del Objetivo General y Validación de Hipótesis | 87 |
| 5.1.2. Contraste con el Estado del Arte | 87 |
| 5.1.3. Impacto y Relevancia en el Contexto Científico-Tecnológico | 88 |
| 5.1.4. Sensibilidad Temporal y Variabilidad Intra-sujeto | 89 |
| 5.2. Conclusiones | 90 |
| 5.2.1. Sobre el Cumplimiento de Objetivos Específicos | 90 |

| | |
|---|------------|
| 5.2.2. Conclusión General | 92 |
| 5.3. Limitaciones del Estudio | 93 |
| 5.3.1. Limitaciones Técnicas | 93 |
| 5.3.2. Limitaciones Metodológicas | 93 |
| 5.3.3. Limitaciones Teóricas | 94 |
| 5.4. Recomendaciones y Trabajos Futuros | 95 |
| 5.4.1. Mejoras en el Hardware y Configuración Experimental | 95 |
| 5.4.2. Propuesta Teórica Avanzada: Métrica Mejorada de Apertura Ocular (Improved EAR) | 95 |
| 5.4.3. Integración con Otras Tecnologías Emergentes | 97 |
| 5.4.4. Nuevas Líneas de Investigación: Salud Ocupacional | 98 |
| 5.5. Reflexión Final | 98 |
| Referencias Bibliográficas | 99 |
| Anexo: Compendio de Formulación Matemática | 103 |

Índice de figuras

| | | |
|------|---|----|
| 2.1. | Imagen descriptiva del ojo humano. | 7 |
| 2.2. | Método de pupila oscura y vector de dirección de la mirada. La pupila aparece como una región oscura debido a la absorción de luz infrarroja por la retina. . . . | 12 |
| 2.3. | Arquitectura del clasificador Random Forest. El diagrama ilustra el mecanismo de ensamble: múltiples árboles de decisión ($Tree_1 \dots Tree_n$) procesan la entrada independientemente. La clase final se determina mediante la mayoría de votos (Majority Voting), lo que reduce la varianza y mejora la generalización del modelo. . . . | 19 |
| 3.1. | Estación de trabajo utilizada para el procesamiento de datos y entrenamiento de la red neuronal. El equipo operó bajo entorno Linux (Ubuntu 24.04 LTS) para maximizar la eficiencia en la gestión de procesos en tiempo real. | 23 |
| 3.2. | Módulo de cámara con sensor GalaxyCore GC0308 adaptado para visión infrarroja cercana (NIR). Se observa la lente con montura M12 ajustable. | 24 |
| 3.3. | Diagrama de flujo de la arquitectura de procesamiento de datos. Se ilustra la secuencia desde la captura de video hasta la generación de resultados biométricos. . . . | 29 |
| 3.4. | Muestra representativa de los ojos capturados durante la sesión experimental. Cada imagen corresponde a un participante distinto del estudio ($N = 15$), evidenciando la variabilidad morfológica inter-sujeto en características como el diámetro pupilar y rasgos que el sistema explota como descriptores biométricos primarios. | 31 |
| 3.5. | Esquema de la cuadrícula de estímulos visuales 3×3 . Los números indican el orden secuencial de presentación. Para una resolución de 1920×1080 , la separación horizontal entre estímulos (Δx) es de 640 px y la vertical (Δy) es de 360 px. | 33 |
| 3.6. | Visualización del protocolo de estímulos (Experimento 1). La matriz de 3×3 puntos rojos sobre fondo negro se utiliza para inducir movimientos sacádicos controlados. Los puntos aparecen de forma secuencial con una duración de 2 segundos por posición, cubriendo la totalidad del campo de visión efectivo. . . | 35 |
| 3.7. | Pipeline de procesamiento de imagen implementado. La secuencia muestra: (Arriba) Mejora de la imagen mediante escala de grises y realce de contraste CLAHE. (Abajo) Detección de la ROI con YOLOv8, binarización inversa adaptativa y superposición final del contorno elíptico ajustado sobre la imagen original. . . . | 42 |

| | |
|--|----|
| 3.8. Esquema de la estimación del vector de mirada. (Izquierda) Identificación del contorno pupilar. (Centro) Cálculo del centro del modelo ocular (punto azul) respecto al centroide de la pupila. (Derecha) Proyección del vector de mirada resultante (línea roja) en el espacio 3D. | 42 |
| 3.9. Correspondencia entre la señal visual capturada y su representación computacional. Izquierda: imagen infrarroja del ojo adquirida por la cámara, donde se aprecia la pupila detectada. Derecha: modelo tridimensional resultante, en el que el diámetro pupilar estimado determina la escala del globo ocular y el vector de mirada (<i>gaze vector</i>) que define su orientación en el espacio 3D. Esta transformación constituye el núcleo del pipeline de estimación oculomotora propuesto. | 44 |
| 4.1. Descomposición vectorial de la señal de mirada en un corto periodo de tiempo. Se compara la señal cruda (negro) con la señal filtrada (colores) para las componentes X, Y y Z del vector de mirada. El filtro Savitzky-Golay ($w = 21$, $p = 3$) elimina el ruido de alta frecuencia manteniendo la fidelidad de los cambios de posición bruscos (sacádicos). | 52 |
| 4.2. Perfil temporal de la velocidad del vector de mirada para un voluntario representativo, obtenido tras la aplicación del filtro Savitzky-Golay. Los picos en rojo identifican eventos sacádicos y microsacádicos detectados mediante el umbral $v > 80^\circ/s$ [26], y las franjas verdes delimitan los períodos de fijación estable entre movimientos balísticos consecutivos. Esta clasificación constituye la base para la extracción de las características biométricas analizadas en las secciones siguientes. | 53 |
| 4.3. Diagrama de dispersión de la Secuencia Principal para la población completa ($N = 15$). Los puntos representan los eventos oculares individuales detectados, mientras que la línea continua indica el ajuste del modelo exponencial de Bahill. La clara adherencia a la curva ($R^2 = 0.9216$) confirma que el sistema captura fielmente la saturación biomecánica del movimiento ocular humano. | 55 |
| 4.4. Perfil cinemático detallado de un sacádico horizontal. Azul (Eje Izq): Velocidad angular mostrando el perfil de campana esperado en un movimiento de aceleración y desaceleración Naranja (Eje Der): La señal de Jerk se mantiene acotada dentro de rangos fisiológicos, sin picos de ruido de alta frecuencia, lo que indica una reconstrucción estable de la trayectoria. | 59 |

| | |
|---|----|
| 4.5. Ranking de importancia de Gini de las características biométricas. Las barras representan la fracción del poder discriminativo total aportada por cada variable al ensemble de 300 árboles. Se observa un predominio de los descriptores morfológicos (Pupil_Mean) sobre los cinemáticos, con una contribución secundaria pero significativa de Main_Seq_Slope y Jerk_Max, validando el carácter híbrido del sistema. | 61 |
| 4.6. Comparación de perfiles biométricos para tres participantes distintos. (Izquierda) El sujeto 1 muestra un perfil orientado a la dinámica (alta velocidad y tasa sacádica). (Centro) El sujeto 2 se distingue por características anatómicas dominantes (mayor tamaño pupilar) y menor reactividad dinámica. (Derecha) El sujeto 3 presenta un perfil balanceado con alta complejidad fractal. Estas "firmas visuales" validan la hipótesis de unicidad del patrón oculomotor. | 63 |
| 4.7. Proyección en 2D (LD1 vs LD2) | 64 |
| 4.8. Proyección en 3D (LD1, LD2, LD3) | 64 |
| 4.9. Espacio de características transformado mediante LDA. Cada color (P1-P14) representa a un participante distinto. Se observa la formación de clústeres compactos y bien definidos, lo que confirma visualmente la separabilidad lineal de las identidades biométricas. | 64 |
| 4.10. Gráfico de sedimentación del análisis de componentes principales. Las barras representan la varianza explicada por cada componente individual, mientras que la línea roja muestra la varianza acumulada. Se requieren 10 componentes para alcanzar el 95 % de la varianza total, lo que evidencia la alta dimensionalidad intrínseca del espacio de características biométricas y la imposibilidad de reducirlo drásticamente sin pérdida significativa de información. | 66 |
| 4.11. Proyección de los datos en las dos primeras componentes principales (PC1 y PC2), que en conjunto explican aproximadamente el 20 % de la varianza total. Se observa un fuerte solapamiento entre las clases, lo que evidencia la limitada separabilidad lineal del espacio biométrico original. | 67 |
| 4.12. Proyección tridimensional de los datos sobre las tres primeras componentes principales (PC1, PC2 y PC3). Aunque se incorpora una dimensión adicional, el solapamiento entre sujetos persiste, confirmando que la varianza explicada por estas componentes es insuficiente para una discriminación fiable. | 67 |
| 4.13. Análisis de componentes principales (PCA) del espacio de características biométricas. Las proyecciones en 2D y 3D muestran la dificultad de separar linealmente a los 15 participantes, lo que justifica la necesidad de métodos no lineales o supervisados. | 67 |

| | |
|---|----|
| 4.14. Visualización mediante t-SNE (perplejidad = 30) de las características biométricas. | 68 |
| 4.15. Proyección mediante UMAP (n_neighbors = 15, min_dist = 0.1) del mismo espacio de características. | 68 |
| 4.16. Visualización no lineal del espacio biométrico mediante t-SNE (izquierda) y UMAP (derecha). Ambas técnicas confirman la existencia de una estructura subyacente, pero al carecer de capacidad de generalización, su utilidad se limita a la exploración visual. Estos resultados refuerzan la necesidad de métodos supervisados como Random Forest para la identificación de sujetos. | 68 |
| 4.17. Matriz de confusión normalizada — <i>Random Forest</i> con $k = 5$ sujetos (exactitud: 92.92 %, confianza media en aciertos: 0.821). La diagonal prácticamente unitaria refleja la alta separabilidad del espacio de características cuando el número de clases es reducido. | 73 |
| 4.18. Matriz de confusión normalizada — <i>Random Forest</i> con $k = 10$ sujetos (exactitud: 92.29 %, confianza media en aciertos: 0.669). La exactitud se mantiene estable respecto a $k = 5$ a pesar de duplicar el número de clases, evidenciando la robustez del sistema en el régimen de escalabilidad moderada. | 74 |
| 4.19. Matriz de confusión normalizada — <i>Random Forest</i> con $k = 15$ sujetos (exactitud: 82.59 %, $\bar{F}1 = 0.823 \pm 0.087$). La diagonal dominante confirma el alto rendimiento global del sistema. Las confusiones fuera de la diagonal se concentran en P6 y P9, sujetos cuyo perfil biométrico presenta mayor solapamiento inter-clase. | 75 |
| 4.20. Distribución espacial de las coordenadas oculares proyectadas mediante un mapeo estrictamente lineal. La ausencia de corrección homográfica evidencia el error de paralaje y la deformación geométrica entre el plano de la cámara y el plano de la pantalla. | 77 |
| 4.21. Mapa de calor de la mirada corregida mediante la matriz de homografía para la cuadrícula 3×3 . Las zonas rojas (alta densidad) coinciden con la ubicación de los 9 puntos de calibración, demostrando que el usuario mantuvo fijaciones estables sobre los objetivos. Las líneas tenues entre clústeres representan las trayectorias sacádicas de transición. | 78 |
| 4.22. Mapa de calor de la mirada para la cuadrícula de 5 puntos. Los 5 clústeres de densidad coinciden con las posiciones objetivo. La menor densidad de puntos de referencia respecto al experimento de 9 puntos se traduce en una menor restricción geométrica de la matriz de homografía, lo que incrementa el error de proyección en las regiones intermedias de la pantalla. | 79 |

| | |
|--|----|
| 4.23. Superposición de la trayectoria espiral teórica (curva de referencia) y la trayectoria de mirada registrada por el sistema. La adherencia de la mirada a la curva teórica evalúa la capacidad del sistema para reconstruir movimientos oculares continuos de seguimiento suave (<i>smooth pursuit</i>), en contraste con las sacadas discretas de los experimentos anteriores. | 80 |
| 4.24. Comparativa de robustez temporal para el sujeto P11. Izquierda (Mañana) : la densidad de aciertos se concentra en la diagonal principal, indicando alta precisión. Derecha (Tarde) : el desplazamiento sistemático de predicciones hacia la columna de P5 refleja la convergencia biométrica selectiva inducida por el cambio de iluminación. | 82 |
| 5.1. Comparación geométrica entre formulaciones. (A) EAR Tradicional utiliza solo 6 puntos de referencia. (B) EAR Mejorado emplea un mallado denso con N líneas de medición, proporcionando mayor robustez. | 96 |

Índice de tablas

| | |
|--|----|
| 3.1. Descripción funcional de los módulos de software del sistema. | 30 |
| 3.2. Caracterización demográfica y condiciones ambientales por participante con incertidumbres asociadas. | 32 |
| 4.1. Métricas finales de rendimiento del modelo YOLOv8n tras 80 épocas de entrenamiento. | 54 |
| 4.2. Estadísticas descriptivas del corpus de sacádicos utilizado para el ajuste de la <i>Main Sequence</i> ($n = 2,157$ eventos, $N = 15$ sujetos). | 56 |
| 4.3. Comparación de los parámetros de la <i>Main Sequence</i> obtenidos experimentalmente frente a los rangos normativos reportados para adultos sanos. | 56 |
| 4.4. Taxonomía de métricas retenidas por el clasificador <i>Random Forest</i> y su rol en la discriminación biométrica. Solo se incluyen las variables con importancia de Gini estadísticamente superior a cero tras el entrenamiento con $ \mathcal{T} = 300$ árboles. | 62 |
| 4.5. Métricas de rendimiento de los algoritmos de clustering no supervisado. | 69 |
| 4.6. Métricas de rendimiento de los clasificadores biométricos en el conjunto de prueba ($k = 15$ sujetos, $n_{test} = 672$). Para el <i>Random Forest</i> , los valores de precisión, recall y F1 se reportan como media \pm desviación estándar entre clases. | 69 |
| 4.7. Hiperparámetros del clasificador <i>Random Forest</i> optimizados mediante búsqueda en rejilla (<i>grid search</i>). | 70 |
| 4.8. Rendimiento del clasificador <i>Random Forest</i> en función del número de sujetos (k). La confianza reportada corresponde a la probabilidad de clase promedio asignada por el ensemble de 300 árboles. | 71 |
| 4.9. Métricas de clasificación por sujeto para el modelo <i>Random Forest</i> ($k = 15$). Precisión: fracción de predicciones de la clase que son correctas. Recall: fracción de instancias reales de la clase correctamente identificadas. F1: media armónica de ambas. | 72 |
| 4.10. Comparativa de precisión espacial entre los tres protocolos de validación del sistema de control de cursor. | 81 |
| 4.11. Análisis estadístico de la deriva biométrica para ocho métricas. Prueba de Mann-Whitney-U (M→T: Mañana vs Tarde). d : d de Cohen. λ : coeficiente de convergencia hacia P5 (Ec. 4.16). En negrita: efectos de mayor relevancia clínica o clasificatoria. | 83 |

| | |
|--|----|
| 5.1. Comparación cuantitativa entre EAR Tradicional y EAR Propuesto. | 96 |
|--|----|

Introducción

El análisis del movimiento ocular permite comprender aspectos fundamentales de la fisiología humana y desarrollar nuevas interfaces tecnológicas. El ojo proporciona información biológica y cinemática única: el patrón del iris y la pupila constituyen identificadores morfológicos robustos, mientras que las trayectorias oculares permiten inferir estados neurológicos y patrones de comportamiento Komogortsev et al. [1].

El desarrollo científico en este campo se ha visto limitado por la barrera económica de los dispositivos comerciales de seguimiento ocular (*eye-trackers*), cuyos costos restringen su aplicación masiva. Esta investigación desarrolla un sistema integral de seguimiento ocular de bajo costo fundamentado en técnicas de procesamiento de imágenes y visión artificial.

El objetivo es doble: validar la viabilidad de la biometría oculomotora para identificación de individuos y evaluar su aplicabilidad como interfaz de control humano-máquina. Se diseñó un sistema de captura basado en sensores CMOS operando a 120 FPS en el espectro infrarrojo cercano (NIR), permitiendo la reconstrucción de la dinámica ocular con alta fidelidad temporal.

La metodología integra principios de óptica geométrica con algoritmos de aprendizaje profundo. Se implementó una arquitectura basada en la red neuronal YOLOv8 para la detección de la pupila, complementada con filtros Savitzky-Golay para el cálculo de derivadas cinemáticas (velocidad, aceleración y *Jerk*). A diferencia de los sistemas biométricos estáticos tradicionales, esta investigación utiliza una huella oculomotora dinámica, demostrando que la interacción entre anatomía ocular y micro-estrategia de movimiento es única para cada sujeto.

Los resultados experimentales, obtenidos de 15 participantes, validan el cumplimiento de leyes fisiológicas fundamentales como la *Main Sequence* Bahill et al. [2] ($R^2 > 0.90$), y demuestran una exactitud de clasificación biométrica del 83.5 % mediante *Random Forest* Breiman [3]. Asimismo, se presenta una prueba de concepto para control de cursor mediante la mirada, estableciendo las bases para tecnologías asistivas accesibles.

Capítulo 1

Planteamiento del Problema

1.1. El problema

Esta investigación se fundamenta en la premisa de que el ojo proporciona información biológica y cinemática única. Morfológicamente, el patrón del iris y las dimensiones de la pupila actúan como identificadores biométricos robustos e irrepetibles. Dinámicamente, el análisis de los movimientos oculares, tales como vibraciones o anomalías en la trayectoria, permite inferir datos significativos sobre el estado fisiológico y neurológico del individuo.

Un aspecto tecnológicamente relevante de este campo es la capacidad de traducir la mirada en comandos de control para la interacción humano-máquina. Estas aplicaciones son críticas en medicina, ofreciendo soluciones de asistencia para personas con movilidad reducida, y en áreas como la neurociencia, interesadas en modelar el comportamiento humano.

El objetivo central es implementar un sistema integral de captura y procesamiento de patrones oculares. El sistema permitirá validar la identificación biométrica y abrir la posibilidad a la creación de herramientas de control por mirada que mejoren la calidad de vida de usuarios con discapacidades motoras.

Díaz y Rojas [4] sostienen que:

El sentido más importante que proporciona la mayor cantidad de información, es el sentido de la vista y por esto, el estudio de la actividad visual mediante la técnica del seguimiento ocular (eye-tracking), se ha convertido en un tema de interés durante los últimos años, para las investigaciones en educación matemática, que incorporan los avances de la neurociencia cognitiva (p. 39).

La cita hace alusión especial a la educación en matemática, ya que las investigadoras utilizan el seguimiento ocular para comprender cómo los estudiantes procesan la lectura de problemas. El seguimiento ocular está siendo estudiado en todas las áreas del conocimiento.

La física proporciona herramientas para abordar y resolver problemas prácticos y actuales. Su capacidad para modelar, analizar y comprender fenómenos naturales permite el desarrollo de soluciones que trascienden teorías abstractas y encuentran aplicaciones directas para mejorar la vida de las personas.

Se plantean cuatro interrogantes que guiarán esta investigación:

1. ¿Cómo implementar un sistema de captura de imágenes o videos del ojo?
2. ¿De qué manera se pueden analizar los datos recolectados mediante técnicas de procesamiento de imágenes?
3. ¿Cómo identificar patrones de movimiento ocular a partir de los datos obtenidos?
4. ¿Cómo se puede evaluar el sistema para su aplicación en el control de un cursor?

1.2. Objetivos

Esta investigación se estructura mediante objetivos específicos. Como indican Ramírez et al. [5]: “los objetivos nos permiten dejar en claro cuál va a ser el alcance de nuestro trabajo de investigación, nos indican el punto de llegada, lo que queremos lograr” (p. 56).

Se establece un objetivo general y cuatro objetivos específicos.

1.2.1. Objetivo general

Implementar un sistema de seguimiento ocular a partir de técnicas de procesamiento de imágenes para el análisis de patrones oculares.

1.2.2. Objetivos específicos

1. Implementar un sistema de captura de imágenes o videos del ojo.
2. Analizar los datos recolectados mediante técnicas de procesamiento de imágenes.
3. Identificar patrones de movimiento ocular a partir de los datos obtenidos.
4. Evaluar el sistema para su posible aplicación en el control de un cursor.

1.3. Justificación

Esta investigación integra conocimientos de física y herramientas tecnológicas para abordar problemas reales en el campo de la asistencia médica y la interacción humano-máquina. Los principios de la física, especialmente en áreas como la óptica y el análisis computacional, se vinculan con la medicina para desarrollar alternativas que mejoren la calidad de vida.

El seguimiento ocular permite la comunicación sin necesidad de movimientos físicos o habla, lo que resulta especialmente útil para personas con discapacidades motoras derivadas de lesiones cerebrales, parálisis cerebral o accidentes cerebrovasculares.

La identificación de patrones de movimiento ocular permite desarrollar herramientas asistivas y aporta al avance interdisciplinario de la física aplicada, la tecnología y la medicina.

Como señala Arias [6], en esta sección de la justificación “deben señalarse las razones por las cuales se realiza la investigación y sus posibles aportes desde el punto de vista teórico o práctico” (p. 105).

El análisis de imágenes permite observar fenómenos invisibles a simple vista, impulsando avances tecnológicos y científicos que mejoran la calidad de vida.

1.4. Organización de la tesis

Este trabajo se encuentra estructurado en cinco capítulos:

Capítulo 1: El problema a estudiar. Se plantea el problema de investigación, se establecen los objetivos y se justifica la importancia del estudio.

Capítulo 2: Marco teórico. Se expone una revisión conceptual sobre los temas fundamentales relacionados con el procesamiento de imágenes, la extracción de información a partir de los datos y conceptos fisiológicos relevantes para la investigación, que sirve como sustento para el desarrollo del proyecto.

Capítulo 3: Metodología. Se describe el enfoque metodológico adoptado, las etapas principales del proceso de recolección y análisis de datos, así como las herramientas y técnicas empleadas para el tratamiento de la información y el desarrollo experimental.

Capítulo 4: Resultados. Se presentan los resultados obtenidos a partir del análisis de los datos, acompañados de su respectiva interpretación y discusión preliminar.

Capítulo 5: Discusión y conclusiones. Se realiza una discusión detallada de los resultados en comparación con investigaciones previas, se exponen las conclusiones alcanzadas y se proponen recomendaciones para futuras investigaciones.

Finalmente, se incluyen varios anexos que complementan el trabajo, tales como programas utilizados, abreviaturas, glosario de términos y documentación técnica relevante.

Capítulo 2

Marco Teórico

2.1. Antecedentes teóricos de la investigación

El seguimiento ocular es una técnica que permite registrar y analizar los movimientos de los ojos, proporcionando información valiosa sobre cómo interactúan los individuos con su entorno visual. Esta tecnología ha sido ampliamente utilizada en diversos campos, como la neurociencia, la psicología, la investigación de la experiencia del usuario y el desarrollo de interfaces hombre-máquina.

La física ha contribuido significativamente a este campo, aportando métodos cuantitativos para el análisis del movimiento ocular. Otras áreas del conocimiento también han desarrollado interés por comprender el fenómeno del movimiento ocular y su aplicación en el desarrollo científico.

A continuación se presentan diversos artículos que constituyen la base teórica de la presente investigación. Estos trabajos sustentan la metodología empleada y las técnicas aplicadas.

2.1.1. Eye Aspect Ratio (EAR) para Detección de Somnolencia

El trabajo de Dewi et al. [7] presenta un método innovador para detectar la somnolencia en conductores en tiempo real, basándose en la métrica conocida como Relación de Aspecto del Ojo (EAR), propuesta originalmente por Soukupová y Čech [8].

Este estudio de Soukupová y Čech es considerado fundamental en la literatura de visión artificial, ya que marcó un cambio de paradigma en el análisis facial: abandonó el procesamiento intensivo de texturas y píxeles en favor de un modelo puramente geométrico basado en puntos de referencia (*landmarks*). Esta abstracción permitió, por primera vez, ejecutar algoritmos de detección de parpadeo con alta precisión en hardware de recursos limitados (como sistemas embebidos), estableciendo el estándar de eficiencia que domina el campo hasta la actualidad.

La técnica se fundamenta en la identificación de 6 puntos de referencia faciales alrededor del contorno ocular. Matemáticamente, el EAR se calcula mediante la relación entre las distancias verticales y la distancia horizontal del ojo, tal como se expresa en la Ecuación 2.1:

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2 \cdot ||p_1 - p_4||} \quad (2.1)$$

Donde p_1 y p_4 corresponden a las comisuras (extremos horizontales), mientras que los pares (p_2, p_6) y (p_3, p_5) representan los puntos del párpado superior e inferior respectivamente. El numerador calcula el promedio de la apertura vertical, y el denominador normaliza esta medida respecto al ancho del ojo, haciendo la métrica invariante a la escala de la imagen y a la distancia del usuario.

En su estudio aplicado, Dewi et al. [7] concluyen que un umbral de EAR de 0.18 ofrece el mejor compromiso entre precisión y rendimiento. Este método tiene un alto potencial para sistemas de seguridad vial en tiempo real, aunque, como se discutirá en el Capítulo 5, la fórmula original presenta limitaciones de estabilidad ante ruido de alta frecuencia que esta tesis busca mitigar mediante una propuesta geométrica densa.

2.1.2. Introduction to Eye Tracking: A Hands-On Tutorial

Gao et al. [9] ofrecen una guía accesible sobre los fundamentos del seguimiento ocular. Combina teoría y práctica para facilitar la comprensión de conceptos clave como el funcionamiento de los dispositivos (cámaras infrarrojas), métricas comunes (fijaciones, sacádicos) y su interpretación en contextos como la interacción humano-computadora.

2.1.3. Remote Photoplethysmography

Allado et al. [10] evalúan la precisión de la fotopletismografía remota (rPPG) para medir la frecuencia respiratoria de forma no invasiva utilizando cámaras convencionales. La rPPG demostró una correlación significativa con métodos tradicionales, evitando el uso de sensores físicos.

2.1.4. AI in Medical Imaging Technology

Según estudios recientes [11], la inteligencia artificial está transformando la imagen médica. Destacan el uso de redes neuronales convolucionales (CNN) y redes generativas antagónicas (GAN) para mejorar la precisión en la detección de anomalías.

2.1.5. Abnormal Ocular Movement in Multiple-System Atrophy

Zhou et al. [12] comparan movimientos oculares en pacientes con atrofia multisistémica y enfermedad de Parkinson. Hallaron que las alteraciones en la velocidad y precisión de los sacádi-

cos voluntarios, así como el nistagmo vertical, son marcadores potenciales para el diagnóstico diferencial.

2.1.6. Aplicación de la convolución de matrices

Giménez et al. [13] abordan la convolución de matrices como herramienta fundamental en el procesamiento de imágenes. Explican cómo los kernels de convolución permiten aplicar filtros (suavizado, detección de bordes) esenciales para las CNN.

2.1.7. Analysis of eye-tracking experiments on Tobii T60

Weigle et al. [14] analizan el desempeño del dispositivo Tobii T60. Los resultados muestran que proporciona datos útiles para estudios de visualización, aunque su exactitud puede verse afectada por movimientos de cabeza y condiciones ambientales, sugiriendo la necesidad de una calibración cuidadosa.

La revisión bibliográfica expuesta evidencia el vacío existente en soluciones de bajo costo, nicho que esta investigación pretende ocupar, integrando fundamentos matemáticos, principios físicos y avances computacionales.

2.2. Bases conceptuales

2.2.1. Anatomía del Ojo Humano

El ojo es un órgano que permite percibir la luz, convirtiéndola en señales eléctricas que el cerebro interpreta. Es un sentido clave para la supervivencia humana.

Según Cárdenas [15], las partes principales son:

Esclera: Capa más externa del ojo, de color blanco, formada por tejido conectivo denso.

Córnea: Capa transparente frontal que ayuda a enfocar la luz.

Pupila: Apertura en el centro del iris que regula la entrada de luz.

Iris: Estructura de color que controla el tamaño de la pupila mediante músculos.

Cristalino: Lente flexible detrás de la pupila que enfoca la luz en la retina.

Músculos ciliares: Rodean al cristalino y modifican su grosor para el enfoque.

Vítreo: Sustancia gelatinosa que llena el interior del ojo.

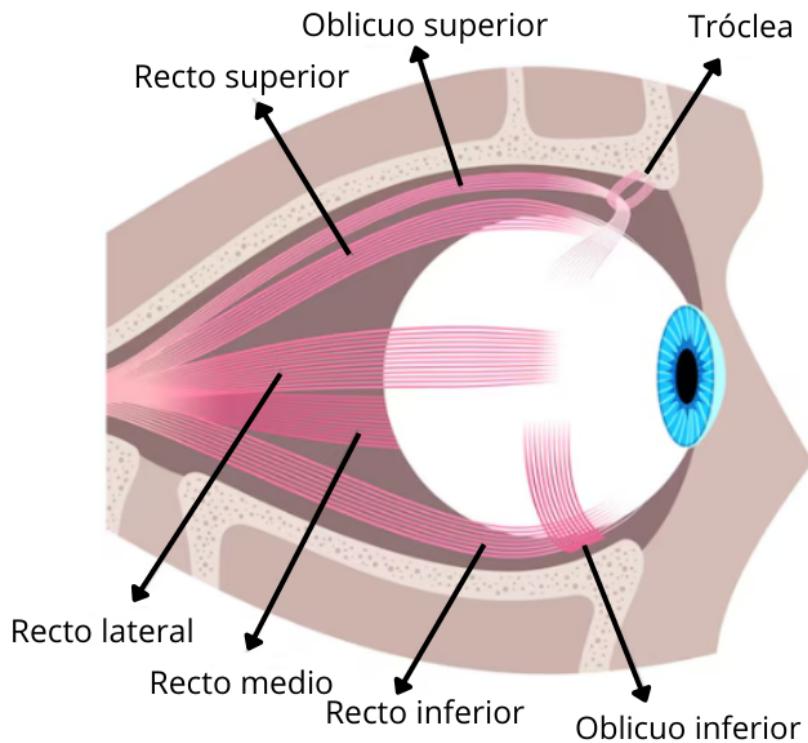


Figura 2.1: Imagen descriptiva del ojo humano.

Retina: Capa sensible a la luz en la parte posterior que contiene fotorreceptores (conos y bastones).

Nervio óptico: Transmite la información visual al cerebro.

2.2.2. Movimientos oculares básicos

El sistema oculomotor permite diversos tipos de movimientos, cada uno con funciones específicas [16]:

Fijaciones

Son períodos durante los cuales el ojo permanece relativamente estático sobre un punto de interés, permitiendo la adquisición de información visual detallada. Típicamente duran entre 200 y 400 ms.

Sacádicos

Movimientos rápidos y balísticos que redirigen la fóvea hacia un nuevo objetivo visual. Son los movimientos más veloces del cuerpo humano, alcanzando velocidades de hasta $900^{\circ}/s$.

Movimientos de seguimiento suave (Smooth Pursuit)

Permiten al ojo seguir un objeto en movimiento, manteniendo su imagen en la fóvea. Su velocidad máxima es de aproximadamente $30^{\circ}/s$.

2.2.3. La secuencia principal (Main Sequence)

Bahill et al. [2] establecieron una relación fundamental conocida como la secuencia principal (*Main Sequence*), que describe la relación matemática y determinista entre la amplitud angular de un movimiento sacádico y su velocidad pico. Esta relación modela el comportamiento balístico del ojo mediante la siguiente expresión exponencial:

$$V_{\max} = V_{\text{sat}} \left(1 - e^{-A/C} \right) \quad (2.2)$$

donde los parámetros tienen la siguiente interpretación física:

- V_{\max} es la velocidad pico alcanzada durante el movimiento (en $^{\circ}/s$).
- A es la amplitud angular total del desplazamiento (en grados).
- V_{sat} es la velocidad de saturación asintótica ($\approx 400\text{-}800^{\circ}/s$ en adultos sanos). Este valor representa el límite biomecánico máximo de contracción de los músculos extraoculares.
- C es la constante de forma ($\approx 10^{\circ}\text{-}20^{\circ}$). Este parámetro define la curvatura de la función y delimita la transición entre movimientos pequeños y grandes.

Desde el punto de vista dinámico, esta ecuación revela dos regímenes de comportamiento del sistema oculomotor:

1. **Régimen Lineal (para $A \ll C$):** En movimientos pequeños (microsacadas o sacádicos cortos), la velocidad crece proporcionalmente a la distancia. Esto implica que la duración del movimiento se mantiene casi constante.
2. **Régimen de Saturación (para $A \gg C$):** A medida que la amplitud aumenta, la velocidad deja de crecer linealmente y se aproxima asintóticamente a V_{sat} . Esto refleja la incapacidad fisiológica de los músculos para acelerar el globo ocular indefinidamente.

Esta ley neurofisiológica es fundamental no solo para describir la cinemática ocular, sino como herramienta de validación de datos: cualquier trayectoria registrada que se desvíe significativamente de esta curva (puntos muy por encima o muy por debajo) puede ser clasificada automáticamente como un artefacto (ruido instrumental, parpadeo) o como una anomalía motora, garantizando así la autenticidad biológica de la señal procesada.

2.3. Sistema de captura y procesamiento

2.3.1. Tecnologías de seguimiento ocular

Los sistemas de seguimiento ocular (eye trackers) emplean diversas tecnologías para registrar la posición y el movimiento de los ojos. La mayoría de los dispositivos modernos utilizan cámaras que operan en el espectro infrarrojo cercano (NIR), específicamente entre 850 nm y 940 nm, longitudes de onda imperceptibles para el ojo humano pero que permiten una captura eficiente en diferentes condiciones de iluminación.

Clasificación según el contacto con el usuario

Dispositivos invasivos: Requieren contacto directo con el ojo, como lentes de contacto con sensores. Ofrecen alta precisión pero son incómodos y no aptos para uso prolongado.

Dispositivos no invasivos: Basados en video-oculografía (VOG), utilizan cámaras externas para rastrear características del ojo. Son los más comunes en investigación y aplicaciones comerciales por su balance entre precisión y comodidad.

Configuraciones de captura

Los sistemas VOG pueden ser:

- **Montados en cabeza (head-mounted):** Cámaras integradas en gafas o cascos, permitiendo movilidad total del usuario. Ideales para estudios en ambientes naturales.
- **Remotos (remote):** Cámaras fijas ubicadas frente al usuario, típicamente integradas en monitores. Requieren que el usuario mantenga la cabeza relativamente estable pero ofrecen mayor comodidad para tareas frente a pantallas.

2.3.2. Fundamentos de la captura por video

La captura de imágenes del ojo se basa en principios de óptica geométrica y procesamiento digital de señales. Una cámara digital registra la luz reflejada por el ojo a una tasa de muestreo

típica de 60 Hz a 1000 Hz, generando secuencias de fotogramas que son procesados en tiempo real o diferido.

Características de las cámaras

Para un sistema de seguimiento ocular efectivo, las cámaras deben cumplir con:

- **Resolución espacial:** Mínimo de 640×480 píxeles para capturar detalles de la pupila y reflexiones corneales.
- **Tasa de muestreo temporal:** Al menos 60 Hz para capturar movimientos básicos; 250-1000 Hz para análisis de sacádicos de alta velocidad.
- **Rango dinámico:** Capacidad de operar bajo diferentes niveles de iluminación ambiente sin saturación.
- **Sincronización:** Los sistemas binoculares requieren sincronización precisa entre cámaras para cálculos estereoscópicos.

2.3.3. Preprocesamiento de la señal

Las imágenes capturadas requieren una serie de transformaciones para extraer información útil sobre la posición ocular. Este pipeline de procesamiento incluye:

Conversión a escala de grises

La mayoría de algoritmos de seguimiento ocular operan sobre imágenes monocromáticas. La conversión se realiza mediante una combinación ponderada de los canales RGB basada en la percepción luminosa humana:

$$I = 0.2989R + 0.5870G + 0.1140B \quad (2.3)$$

donde R, G, B son los componentes de color originales.

Umbralización global

Permite segmentar regiones de interés aplicando un umbral T . La binarización se define como:

$$I'(x,y) = \begin{cases} 255, & \text{si } I(x,y) > T \\ 0, & \text{si } I(x,y) \leq T \end{cases} \quad (2.4)$$

Esto facilita la detección de bordes de la pupila, que típicamente aparece como la región más oscura de la imagen del ojo.

Filtro de Savitzky-Golay

Para el preprocesamiento de las señales oculomotoras, se optó por el filtro digital de Savitzky-Golay en lugar de los métodos tradicionales de promedio móvil. Esta técnica realiza un suavizado basado en el ajuste de mínimos cuadrados locales mediante polinomios de bajo grado dentro de una ventana deslizante. Su operación se describe mediante la convolución discreta:

$$y[i] = \sum_{k=-M}^{M} c_k x[i+k] \quad (2.5)$$

donde $x[i]$ es la señal original, $y[i]$ la señal suavizada, y c_k son los coeficientes de convolución que dependen del orden del polinomio p y del tamaño de la ventana $2M + 1$.

La principal ventaja de este enfoque radica en su capacidad para preservar la forma de la onda (*waveform preservation*). A diferencia de los filtros de media móvil, que actúan como filtros paso bajo agresivos y tienden a atenuar la amplitud de los picos (reduciendo artificialmente la velocidad máxima medida) y ensanchar el ancho temporal de los eventos, el filtro Savitzky-Golay mantiene fielmente los momentos estadísticos de orden superior de la señal. Esto resulta crítico para el análisis de movimientos sacádicos, donde la precisión en la detección de la velocidad pico y la aceleración es prioritaria frente a la simple reducción de ruido estocástico.

Método de la pupila oscura (Dark Pupil)

Esta técnica utiliza iluminación infrarroja fuera del eje óptico de la cámara (off-axis illumination). La retina absorbe la luz infrarroja, resultando en una pupila que aparece oscura en la imagen, mientras que la esclera y el iris reflejan la luz. Este alto contraste facilita la segmentación automática de la pupila mediante algoritmos de detección de contornos o ajuste de elipses.

2.4. Caracterización biométrica y dinámica de la señal

El análisis moderno de movimientos oculares trasciende la simple ubicación de la mirada (dónde se mira) para enfocarse en la dinámica del movimiento (cómo se mira). Esta investigación se fundamenta en la extracción de características matemáticas avanzadas que describen la complejidad, suavidad y respuesta fisiológica del sistema oculomotor, permitiendo la identificación de patrones únicos por sujeto.

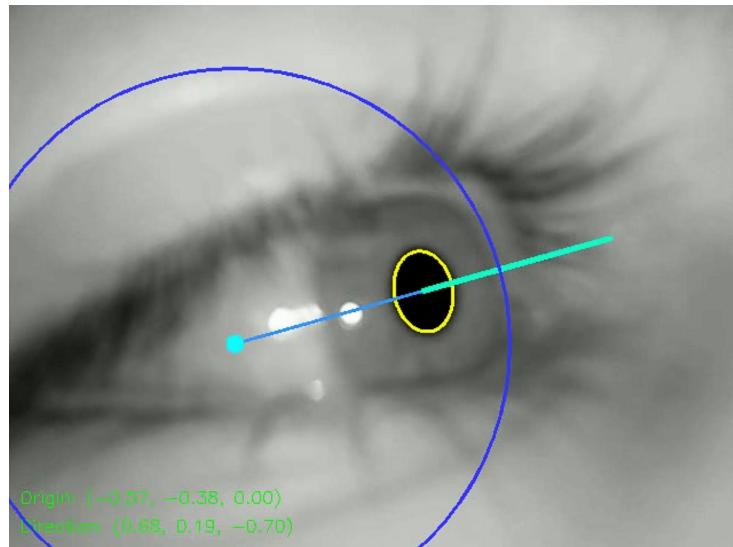


Figura 2.2: Método de pupila oscura y vector de dirección de la mirada. La pupila aparece como una región oscura debido a la absorción de luz infrarroja por la retina.

2.4.1. Biometría oculomotora

La biometría oculomotora se basa en la premisa de que los patrones de movimiento ocular son idiosincrásicos; es decir, únicos para cada individuo. Según Komogortsev et al. [1], esta singularidad surge de la interacción entre dos factores:

- **Factores biomecánicos:** Propiedades anatómicas invariables como la fuerza de los músculos extraoculares, la masa y momento de inercia del globo ocular, la viscosidad de los tejidos orbitales y la elasticidad de los ligamentos suspensorios. Estas características determinan la respuesta mecánica del sistema y son difíciles de replicar entre individuos.
- **Factores neurológicos:** Estrategias cognitivas de procesamiento visual, patrones atencionales, y la eficiencia de las vías neuronales que controlan la planificación y ejecución motora. Estos elementos reflejan diferencias en la conectividad cerebral y los procesos de toma de decisiones visuales.

A diferencia de biometrías estáticas (como la huella dactilar o el iris), la biometría oculomotora es una biometría comportamental que permite una autenticación continua durante la interacción del usuario con un sistema, sin requerir acciones explícitas de verificación.

2.4.2. Pupilometría dinámica y cognitiva

El diámetro pupilar no responde únicamente a los cambios de iluminación ambiental a través del reflejo fotomotor, sino que presenta fluctuaciones oscilatorias vinculadas directamente a la actividad del Sistema Nervioso Autónomo (SNA).

Investigaciones clásicas como [17] establecen que la dilatación pupilar y, particularmente, su velocidad de cambio correlacionan significativamente con:

- La carga cognitiva y el esfuerzo mental requerido por una tarea
- El estado emocional y niveles de excitación
- Procesos de toma de decisiones y resolución de problemas

En este estudio, se analiza la velocidad máxima de cambio del diámetro pupilar (*Pupil Velocity*, medida en mm/s) como una métrica de la reactividad del sistema nervioso del individuo. Esta característica temporal, sumada al análisis de la amplitud de las fluctuaciones pupilares, aporta información valiosa para la discriminación entre sujetos, especialmente cuando se combina con características de los movimientos sacádicos.

2.4.3. Análisis de la calidad del movimiento: Jerk y suavidad

Para evaluar la eficiencia y fluidez del control motor ocular, se emplea el concepto físico de *Jerk* (también llamado sobreaceleración o tirón), definido formalmente como la derivada temporal de la aceleración, o equivalentemente, la tercera derivada de la posición respecto al tiempo:

$$J(t) = \frac{d^3x(t)}{dt^3} = \frac{da(t)}{dt} \quad (2.6)$$

donde $x(t)$ es la posición angular del ojo en el tiempo t , y $a(t)$ es la aceleración.

El sistema nervioso humano optimiza los movimientos voluntarios para minimizar el Jerk, logrando trayectorias suaves y energéticamente eficientes. Un índice elevado de Jerk o una baja suavidad (*smoothness*) en la transición de velocidades durante un sacádico puede indicar:

- Fatiga muscular o neurológica
- Patologías del sistema oculomotor o cerebeloso
- Características motoras específicas de un individuo que no posee un control motor fino
- Efectos de sustancias psicoactivas o medicamentos

Para cuantificar la suavidad, se utiliza comúnmente el logaritmo de la magnitud adimensional del Jerk (Log Dimensionless Jerk, LDLJ), que permite comparaciones entre movimientos de diferentes amplitudes y duraciones.

2.4.4. Análisis de señales no lineales y complejidad fractal

Los movimientos oculares exhiben comportamientos que no pueden ser descritos completamente por la estadística lineal tradicional (media, varianza, correlación), presentando estructuras caóticas y propiedades fractales que reflejan la complejidad del sistema de control neuromuscular.

Para cuantificar esta complejidad intrínseca, se utiliza la Dimensión Fractal de Higuchi (Higuchi Fractal Dimension, HFD).

Propuesto por Higuchi [18], este algoritmo calcula la rugosidad o complejidad geométrica de una serie temporal sin requerir que la señal sea estacionaria. El método construye múltiples subsecuencias de la señal original con diferentes intervalos de tiempo y calcula la longitud de curva de cada una.

La dimensión fractal resultante (D_H) toma valores en el rango [1, 2]:

- $D_H \approx 1$: Señal suave y altamente predecible (comportamiento casi lineal)
- $D_H \approx 2$: Señal altamente irregular y compleja (similar al ruido browniano)

En el contexto de movimientos oculares, la dimensión fractal permite distinguir entre:

- Una exploración visual fluida y sistemática (baja dimensión fractal): característica de usuarios con estrategias de búsqueda eficientes y planificadas
- Una exploración errática y compleja (alta dimensión fractal): indicativa de búsqueda no estructurada, incertidumbre visual o déficits atencionales

Esta métrica captura aspectos de la estrategia cognitiva de búsqueda visual de cada participante que son complementarios a las métricas clásicas de velocidad y precisión.

2.5. Técnicas de análisis y clasificación

Una vez extraídas las características biométricas y dinámicas de las señales oculares, es necesario emplear técnicas estadísticas y de aprendizaje automático para analizar, reducir la dimensionalidad y clasificar los patrones de movimiento ocular con el objetivo de identificar individuos o detectar estados fisiológicos.

2.5.1. Estadística descriptiva

El análisis de los datos oculométricos requiere una caracterización inicial mediante estadística descriptiva. Estas medidas permiten resumir el comportamiento de las variables registradas y evaluar la calidad de la señal adquirida.

Media aritmética (μ): Se utiliza como medida de tendencia central para determinar valores representativos, tales como la duración promedio de una fijación o la velocidad media de un movimiento sacádico. Matemáticamente se define como:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.7)$$

donde n es el número de observaciones y x_i es cada valor individual.

Desviación estándar (σ): Es una métrica fundamental para evaluar la precisión del sistema de seguimiento ocular. En este contexto, la precisión se refiere a la variabilidad o dispersión de los datos adquiridos bajo condiciones repetidas. Se calcula como:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (2.8)$$

Una desviación estándar baja indica alta consistencia entre mediciones repetidas bajo las mismas condiciones, lo cual es indicativo de un sistema estable y confiable. Valores típicos de precisión en sistemas comerciales de seguimiento ocular remoto son de 0.5° a 1° de ángulo visual.

2.5.2. Análisis Discriminante Lineal (LDA)

El Análisis Discriminante Lineal (Linear Discriminant Analysis, LDA) es una técnica estadística de aprendizaje supervisado utilizada para la reducción de dimensionalidad y la clasificación de patrones. Su propósito es proyectar un conjunto de datos de alta dimensión (d características) a un subespacio de menor dimensión (k dimensiones, donde $k < d$), maximizando la separabilidad lineal entre las distintas clases.

A diferencia del Análisis de Componentes Principales (PCA), que es un método no supervisado centrado en maximizar la varianza total de los datos sin considerar las etiquetas de clase, el LDA busca explícitamente modelar la diferencia entre las clases. Para lograr esto, el algoritmo maximiza el Criterio de Fisher, una función objetivo que busca simultáneamente dos metas:

1. **Minimizar la varianza dentro de cada clase** (*within-class scatter, S_W*): Asegura que las muestras de una misma clase se agrupen de forma compacta en el espacio proyectado.

2. **Maximizar la distancia entre las medias de las diferentes clases** (*between-class scatter, S_B*): Garantiza que las clases estén lo más separadas posible.

Matemáticamente, el criterio de Fisher se expresa como:

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (2.9)$$

donde w es el vector de proyección que se busca optimizar.

El resultado es una transformación lineal que preserva la información discriminatoria más relevante, facilitando así la construcción de clasificadores más eficientes y robustos ante la variabilidad de los datos. En el contexto de biometría ocular, el LDA permite reducir un conjunto de decenas de características (velocidades, aceleraciones, dimensiones fractales, etc.) a un espacio de pocas dimensiones donde las diferencias entre sujetos son más evidentes.

2.5.3. Máquinas de Vectores de Soporte (SVM)

Las Máquinas de Vectores de Soporte (Support Vector Machines, SVM) constituyen un conjunto de algoritmos de aprendizaje supervisado desarrollados por Vapnik [19], fundamentados en la teoría del aprendizaje estadístico y la minimización del riesgo estructural.

El objetivo central de una SVM es encontrar el hiperplano óptimo de separación en un espacio multidimensional. Este hiperplano se elige de tal manera que maximice el margen entre las clases, entendiendo el margen como la distancia geométrica entre la frontera de decisión y los puntos de entrenamiento más cercanos de cada clase, a los cuales se les denomina vectores de soporte.

Formulación matemática

Para un problema de clasificación binaria, donde los datos de entrenamiento son (x_i, y_i) con $x_i \in \mathbb{R}^d$ y $y_i \in \{-1, +1\}$, el hiperplano de separación se define como:

$$w^T x + b = 0 \quad (2.10)$$

El problema de optimización busca maximizar el margen, lo cual equivale a minimizar $\|w\|^2$ sujeto a las restricciones:

$$y_i(w^T x_i + b) \geq 1, \quad \forall i \quad (2.11)$$

Kernel RBF para datos no lineales

Debido a que los datos biométricos y oculares frecuentemente presentan fronteras de decisión complejas y no lineales, este trabajo implementa SVM utilizando el truco del kernel, específicamente la Función de Base Radial (Radial Basis Function, RBF):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (2.12)$$

donde γ es un hiperparámetro que controla la influencia de cada vector de soporte.

Esta técnica matemática permite mapear implícitamente los vectores de entrada originales a un espacio de características de mayor (incluso infinita) dimensión donde las clases se vuelven linealmente separables. Esto dota al modelo de la capacidad para capturar relaciones no lineales complejas en la dinámica ocular sin incurrir en un costo computacional prohibitivo, ya que el cálculo se realiza mediante productos internos en el espacio original a través de la función kernel.

2.5.4. Random Forest (Bosque Aleatorio)

Random Forest es un algoritmo de aprendizaje conjunto (*ensemble learning*) diseñado para tareas de clasificación y regresión, propuesto originalmente por Breiman [3]. Su funcionamiento se basa en la construcción de una colección de múltiples clasificadores base, específicamente árboles de decisión, durante la etapa de entrenamiento, para luego combinar sus predicciones mediante votación.

Principios fundamentales

La robustez del algoritmo radica en combinar dos técnicas complementarias:

1. **Bagging (Bootstrap Aggregating):** Para construir cada árbol, se genera una muestra de entrenamiento distinta mediante un muestreo con reemplazo (*bootstrap sampling*) del conjunto de datos original. Esto significa que cada árbol se entrena con aproximadamente el 63.2 % de las muestras únicas, mientras que el resto (out-of-bag samples) pueden usarse para validación interna.
2. **Selección aleatoria de características:** En cada nodo de decisión del árbol, el algoritmo no evalúa todas las d variables posibles para determinar la mejor partición, sino únicamente un subconjunto aleatorio de \sqrt{d} características (para clasificación) o $d/3$ (para regresión). Esta aleatorización fuerza a los árboles a explorar diferentes aspectos de los datos.

Criterio de Selección de Atributos: Índice de Gini

Para determinar la división óptima en cada nodo de los árboles que conforman el bosque, el algoritmo utiliza el Índice de Impureza de Gini. Esta métrica mide la probabilidad de clasificar incorrectamente un elemento elegido al azar si se etiquetara de acuerdo con la distribución de clases en el subconjunto de datos.

Matemáticamente, para un nodo dado t , el índice de Gini $G(t)$ se define como:

$$G(t) = 1 - \sum_{i=1}^C p(i|t)^2 \quad (2.13)$$

Donde:

- C es el número total de clases (en este estudio, el número de participantes).
- $p(i|t)$ es la probabilidad relativa de la clase i en el nodo t .

Un índice de Gini igual a 0 indica un nodo puro (todos los elementos pertenecen a la misma clase), mientras que un valor cercano a 1 indica una distribución aleatoria de clases. El algoritmo busca, en cada paso, la característica y el punto de corte que maximicen la reducción de la impureza total, permitiendo así que el modelo aprenda los patrones más discriminativos de las señales oculomotoras.

Importancia de las Variables (Feature Importance)

Más allá de la clasificación, el algoritmo de Random Forest permite cuantificar la relevancia de cada métrica extraída (como la velocidad sacádica o la dimensión fractal) mediante la *Importancia de Características por Impureza de Gini*.

Este valor se calcula sumando todas las reducciones del índice de Gini que ocurren cada vez que una variable específica se utiliza para dividir un nodo, promediando dicho valor sobre todos los árboles del bosque. Matemáticamente, la importancia de una característica X_j se define como:

$$I(X_j) = \frac{1}{T} \sum_{t=1}^T \sum_{n \in nodes(t), v(n)=X_j} \Delta G(n) \quad (2.14)$$

Donde:

- T es el número total de árboles.
- $\Delta G(n)$ es el decremento en la impureza de Gini al realizar la división en el nodo n .
- $v(n) = X_j$ indica que la variable X_j fue la utilizada para realizar la partición en ese nodo.

Aquellas variables que generen nodos más puros (reducciones de Gini más drásticas) de manera consistente en todo el bosque obtendrán una puntuación de importancia mayor. En el contexto de esta tesis, este análisis permite identificar cuáles de los parámetros oculomotores son biométricamente más significativos para distinguir la identidad de un individuo frente a la población de estudio.

Predicción y ventajas

La clasificación de una nueva instancia se obtiene mediante votación por mayoría simple de todos los árboles del bosque:

$$\hat{y} = \text{moda}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (2.15)$$

donde $h_t(x)$ es la predicción del árbol t -ésimo y T es el número total de árboles.

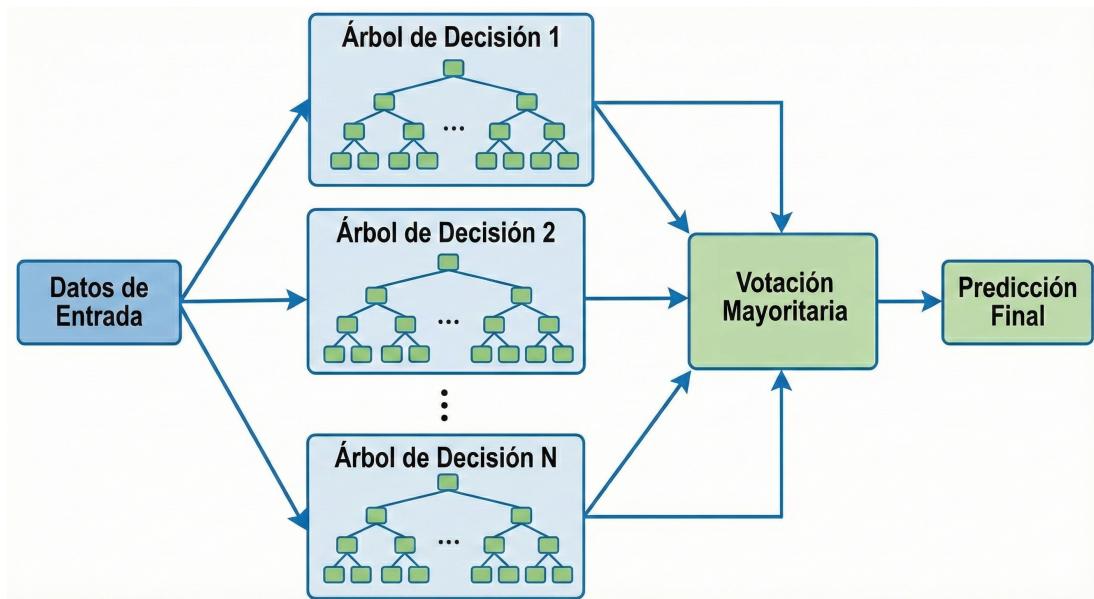


Figura 2.3: Arquitectura del clasificador Random Forest. El diagrama ilustra el mecanismo de ensamble: múltiples árboles de decisión ($Tree_1 \dots Tree_n$) procesan la entrada independientemente. La clase final se determina mediante la mayoría de votos (Majority Voting), lo que reduce la varianza y mejora la generalización del modelo.

Como se ilustra en la Figura 2.3, esta estrategia de doble aleatorización reduce significativamente la correlación entre los árboles individuales y mitiga eficazmente el riesgo de sobreajuste (*overfitting*), resultando en un modelo con una alta capacidad de generalización. Adicionalmente, Random Forest proporciona estimaciones automáticas de la importancia de cada característica, lo cual es valioso para interpretar qué aspectos de los movimientos oculares son más discriminativos para la identificación de individuos.

Capítulo 3

Marco Metodológico

Este capítulo describe el marco metodológico de la investigación: el diseño experimental, el sistema de captura implementado y las técnicas de procesamiento de señales utilizadas.

3.1. Diseño y Tipo de Investigación

3.1.1. Enfoque Cuantitativo

La presente investigación adopta un enfoque cuantitativo, sustentado en la recolección sistemática de datos numéricos y su posterior análisis mediante técnicas estadísticas y computacionales. Esta elección metodológica se justifica por la naturaleza física y medible de las variables bajo estudio: posiciones espaciales del ojo expresadas en coordenadas (x, y), velocidades angulares (en $^{\circ}/s$), aceleraciones, y otras métricas cinemáticas que describen la dinámica del sistema oculomotor. El uso de variables numéricas permite cuantificar de manera precisa los fenómenos físicos involucrados en el movimiento ocular. Según Ramírez et al. [5], el enfoque cuantitativo es apropiado cuando se busca establecer patrones de comportamiento, probar hipótesis y generalizar resultados a partir de muestras representativas mediante el análisis estadístico. Las trayectorias oculares registradas se traducen en series temporales multidimensionales que pueden ser analizadas mediante herramientas matemáticas rigurosas. Las características extraídas —tales como la velocidad pico de los sacádicos, el índice de suavidad (Jerk), la dimensión fractal de Higuchi, y las variaciones del diámetro pupilar— constituyen descriptores cuantitativos que permiten:

- Caracterizar objetivamente el comportamiento oculomotor de cada participante
- Comparar estadísticamente las diferencias individuales
- Aplicar algoritmos de aprendizaje automático para la clasificación de patrones
- Validar la consistencia de las mediciones mediante análisis de precisión y reproducibilidad

Este enfoque sigue los principios de la física experimental, donde los fenómenos naturales se describen mediante modelos matemáticos y se validan a través de la replicabilidad y el análisis cuantitativo de los resultados obtenidos.

3.1.2. Diseño Experimental

El estudio se enmarca dentro de un diseño experimental de tipo controlado, en el cual se manipulan deliberadamente las variables independientes para observar su efecto sobre las variables dependientes del sistema oculomotor. Específicamente, se diseñaron estímulos visuales estandarizados que permiten registrar y analizar las respuestas oculares bajo condiciones replicables.

Variables del estudio

El diseño experimental se estructura en torno a las siguientes variables:

Variables independientes (manipuladas):

- *Posición espacial del estímulo visual:* Una cuadrícula de 3×3 puntos que cubren el campo visual de la pantalla, generando 9 posiciones discretas que provocan movimientos sacádicos dirigidos.
- *Duración de presentación del estímulo:* Cada punto permanece visible durante 2 segundos, tiempo suficiente para garantizar la estabilización de la fijación ocular según los criterios establecidos por Duchowski [16].
- *Secuencia de presentación:* Los estímulos se presentan siguiendo un orden determinado (de izquierda a derecha, de arriba hacia abajo), lo que induce un patrón predecible de movimientos oculares.

Variables dependientes (observadas):

- Coordenadas espaciales del centroide de la pupila $(x_p(t), y_p(t))$ en función del tiempo
- Velocidad angular del movimiento ocular $v(t)$ [$^{\circ}/s$]
- Aceleración angular $a(t)$ [$^{\circ}/s^2$]
- Índice de suavidad del movimiento (Jerk) $J(t)$ [$^{\circ}/s^3$]
- Diámetro pupilar $d(t)$ [mm] y su tasa de cambio
- Duración de las fijaciones y características de los movimientos sacádicos

Variables controladas:

- Distancia sujeto-pantalla: 60 ± 1 cm
- Resolución y frecuencia de muestreo de la cámara

- Características físicas del estímulo (tamaño, contraste, color de fondo)

Este diseño permite establecer relaciones causales entre los estímulos presentados y las respuestas oculomotoras registradas, asegurando la validez interna del experimento mediante el control riguroso de factores que podrían introducir variabilidad no deseada en las mediciones.

3.2. Sistema Experimental y Herramientas

3.2.1. Hardware e Instrumentación

El sistema experimental desarrollado integra componentes de hardware que permiten la captura, procesamiento y almacenamiento de datos oculométricos en tiempo real.

Equipo de cómputo

El procesamiento de las señales oculares requiere capacidad computacional suficiente para ejecutar algoritmos de visión por computadora y aprendizaje automático en tiempo real. El equipo utilizado presenta las siguientes características:

- **Procesador:** Intel Core i7 de 11^a generación (8 núcleos, 16 hilos)
- **Memoria RAM:** 20 GB DDR4 a 3200 MHz
- **Sistema operativo:** Ubuntu 24.04 LTS (distribución Linux optimizada para desarrollo científico)
- **Almacenamiento:** SSD de 512 GB para lectura/escritura rápida de datos
- **Tarjeta gráfica:** GPU integrada Intel (utilizada para acelerar operaciones de procesamiento de imágenes)

Este equipo proporciona el poder de cómputo necesario para:

1. Procesar flujos de video a 30-120 FPS en tiempo real
2. Ejecutar algoritmos de detección y extracción de características
3. Entrenar y validar modelos de aprendizaje automático
4. Almacenar y gestionar grandes volúmenes de datos experimentales



Figura 3.1: Estación de trabajo utilizada para el procesamiento de datos y entrenamiento de la red neuronal. El equipo operó bajo entorno Linux (Ubuntu 24.04 LTS) para maximizar la eficiencia en la gestión de procesos en tiempo real.

Sistema de captura de video

La captura de imágenes del ojo se realizó mediante un módulo de cámara infrarroja especializado para visión artificial, basado en el sensor GC0308, cuyas especificaciones técnicas se detallan a continuación:

- **Sensor:** GalaxyCore GC0308 (CMOS de 1/6.5 pulgadas)
- **Espectro de operación:** Infrarrojo cercano (NIR) con sensibilidad pico a 850 nm (sin filtro IR-cut)
- **Resolución operativa:** 320×240 píxeles (QVGA) para maximizar la tasa de cuadros
- **Frecuencia de muestreo:** Configurada a 120 FPS constantes
- **Interfaz de datos:** USB 2.0 de alta velocidad
- **Formato de salida:** MJPEG
- **Enfoque:** Lente de enfoque fijo manual (montura M12) ajustado para macro-distancia



Figura 3.2: Módulo de cámara con sensor GalaxyCore GC0308 adaptado para visión infrarroja cercana (NIR). Se observa la lente con montura M12 ajustable.

Se seleccionó una frecuencia de muestreo operativa de 120 Hz fundamentada en el teorema de muestreo de Nyquist-Shannon. Dado que el ancho de banda fisiológico de los movimientos oculares sacádicos convencionales se sitúa en el rango de 20-25 Hz, una tasa de 120 FPS no solo satisface la condición de Nyquist ($120 > 50$), sino que proporciona un factor de sobremuestreo significativo. Esto permite una reconstrucción temporal fidedigna de la dinámica ocular, minimizando el desenfoque de movimiento (*motion blur*) durante las aceleraciones rápidas del ojo.

Configuración geométrica y validez de la captura:

A diferencia de los sistemas remotos, el dispositivo se configuró como un sistema montado en cabeza. La cámara se acopló mecánicamente a la estructura de unos lentes, ubicada específicamente en la parte inferior derecha del campo visual y orientada con un ángulo de elevación hacia el globo ocular derecho. Esta configuración de proximidad (distancia focal < 5 cm) garantiza que la región de interés (ROI) ocupe la totalidad del sensor. A pesar de utilizar una resolución QVGA (320×240), la densidad de píxeles efectiva sobre el área de la pupila es superior a la de una cámara HD ubicada a distancia. La imagen resultante contiene exclusivamente información del ojo, eliminando ruido del entorno y reduciendo la carga computacional en el post-procesamiento, lo que facilita la detección precisa del centroide pupilar.

3.2.2. Entorno y Condiciones de Iluminación

El control de las condiciones ambientales es fundamental para garantizar la reproducibilidad de las mediciones y la calidad de las imágenes capturadas, particularmente cuando se utiliza el método de pupila oscura descrito en el marco teórico (Capítulo 2).

Configuración espacial

La configuración geométrica del sistema se estandarizó para todos los participantes:

- **Distancia sujeto-pantalla:** 60 cm (medida desde la posición promedio de los ojos hasta el centro de la pantalla)
- **Altura de la pantalla:** Ajustada de manera que el centro de la pantalla coincida con la altura de los ojos del participante en posición erguida
- **Ángulo de visión:** 0° (pantalla perpendicular a la línea de visión primaria)
- **Posición de la cámara:** Montada en la parte inferior derecha de los lentes, apuntando aproximadamente a 45° hacia el ojo

Esta distancia de 60 cm fue seleccionada considerando que:

1. Es una distancia ergonómica estándar para el trabajo frente a pantallas según las normas ISO 9241-5
2. Maximiza la resolución angular del sistema de seguimiento sin comprometer la comodidad del participante

Control de iluminación

Para implementar eficazmente el método de pupila oscura y maximizar el contraste entre la pupila y el iris, se establecieron las siguientes condiciones de iluminación:

- **Iluminación ambiental:** Nivel de iluminancia controlado entre 500-900 lux, correspondiente a condiciones de oficina estándar. Todos los datos se recopilaron en este rango de lux, se corroboró usando un luxómetro UNIT UT383.

Método de pupila oscura: Gracias a la sensibilidad espectral del sensor GC0308 en el rango del infrarrojo cercano (sin filtro IR-cut), este estudio implementa el método de pupila oscura aprovechando la iluminación ambiental como fuente pasiva. A diferencia de los sensores que operan exclusivamente en el espectro visible, este dispositivo capta la radiación infrarroja del

entorno, la cual es absorbida eficientemente por la retina pero reflejada por el iris y la esclera. Este fenómeno físico maximiza el contraste natural de la imagen, haciendo que la pupila aparezca significativamente más oscura sin necesidad de emisores activos (LEDs), lo que optimiza la segmentación mediante las técnicas de umbralización detalladas en la sección 3.4.2.

3.2.3. Herramientas de Software

El pipeline de procesamiento, desde la adquisición de imágenes hasta la clasificación de patrones, se implementó íntegramente en Python 3.10, seleccionado por su amplio ecosistema de bibliotecas científicas y su capacidad para prototipado rápido de algoritmos.

Procesamiento de imágenes y visión por computadora

- **OpenCV (cv2) 4.8.0:** Biblioteca fundamental para la captura de video en tiempo real, aplicación de filtros de preprocesamiento (Gaussian, Bilateral), operaciones morfológicas, detección de contornos, y transformaciones geométricas. Esta biblioteca implementa eficientemente los algoritmos descritos en el marco teórico para la segmentación de la pupila.
- **Ultralytics YOLOv8 8.0.x:** Framework de aprendizaje profundo utilizado para la implementación, entrenamiento y validación de la red neuronal convolucional encargada de la detección y seguimiento de la pupila. Su integración permitió inferencias de alta velocidad (tiempo real) con el modelo personalizado.
- **NumPy 1.24.3:** Biblioteca fundamental para operaciones numéricas eficientes sobre matrices multidimensionales. Se utilizó extensivamente para:
 - Manipulación de arrays de imágenes (tensores de dimensión $H \times W \times C$)
 - Cálculo de derivadas numéricas para obtener velocidad, aceleración y Jerk
 - Operaciones de álgebra lineal (productos matriciales, descomposiciones)

Procesamiento de señales

- **SciPy 1.11.2:** Conjunto de herramientas para computación científica que se empleó para:
 - Implementación del filtro Savitzky-Golay (`scipy.signal.savgol_filter`)
 - Cálculo de transformadas de Fourier para análisis frecuencial
 - Interpolación de datos mediante splines cúbicos
 - Optimización de parámetros mediante métodos de mínimos cuadrados

Análisis estadístico y visualización

- **Pandas 2.0.3:** Biblioteca para manipulación y análisis de datos estructurados. Se utilizó para:
 - Organización de las características extraídas en DataFrames
 - Cálculo de estadísticas descriptivas (media, desviación estándar, cuartiles)
 - Exportación de resultados a formatos CSV y Excel
- **Matplotlib 3.7.2 y Seaborn 0.12.2:** Bibliotecas complementarias para visualización de datos. Se emplearon para:
 - Generación de gráficos de trayectorias oculares
 - Visualización de distribuciones de características mediante histogramas y boxplots
 - Creación de matrices de confusión para evaluar los clasificadores
 - Gráficos de la secuencia principal (velocidad vs. amplitud de sacádicos)

Aprendizaje automático

- **Scikit-learn 1.3.0:** Biblioteca integral de aprendizaje automático que proporciona implementaciones eficientes de los algoritmos descritos en el marco teórico:
 - `LinearDiscriminantAnalysis`: Para reducción de dimensionalidad preservando la separabilidad entre clases
 - `SVC`: Máquinas de Vectores de Soporte con kernel RBF para clasificación no lineal
 - `RandomForestClassifier`: Clasificador basado en ensambles de árboles de decisión
 - `StandardScaler`: Normalización de características para evitar sesgos por diferencias de escala
 - `train_test_split`: División estratificada de datos en conjuntos de entrenamiento y prueba
 - Métricas de evaluación: `accuracy_score`, `classification_report`, `confusion_matrix`

3.2.4. Disponibilidad del Código

El código fuente completo del sistema desarrollado, incluyendo los scripts de procesamiento, análisis y visualización, está disponible públicamente en el repositorio de GitHub: <https://github.com/LilkongW/Tesis3D.git>.

Este repositorio contiene:

- Los scripts de captura y procesamiento de imágenes
- El modelo YOLOv8 entrenado para detección pupilar
- Los algoritmos de extracción de métricas biométricas
- Los scripts de análisis estadístico y generación de gráficos

3.2.5. Arquitectura del Flujo de Datos (Pipeline)

El sistema se diseñó bajo una arquitectura modular secuencial compuesta por tres etapas de procesamiento, cada una gestionada por scripts específicos desarrollados en Python. Este flujo garantiza la trazabilidad de los datos desde la captura cruda hasta el análisis estadístico final. La Figura 3.3 ilustra la secuencia de transformación de los datos, vinculando los algoritmos descritos con los módulos de software implementados.

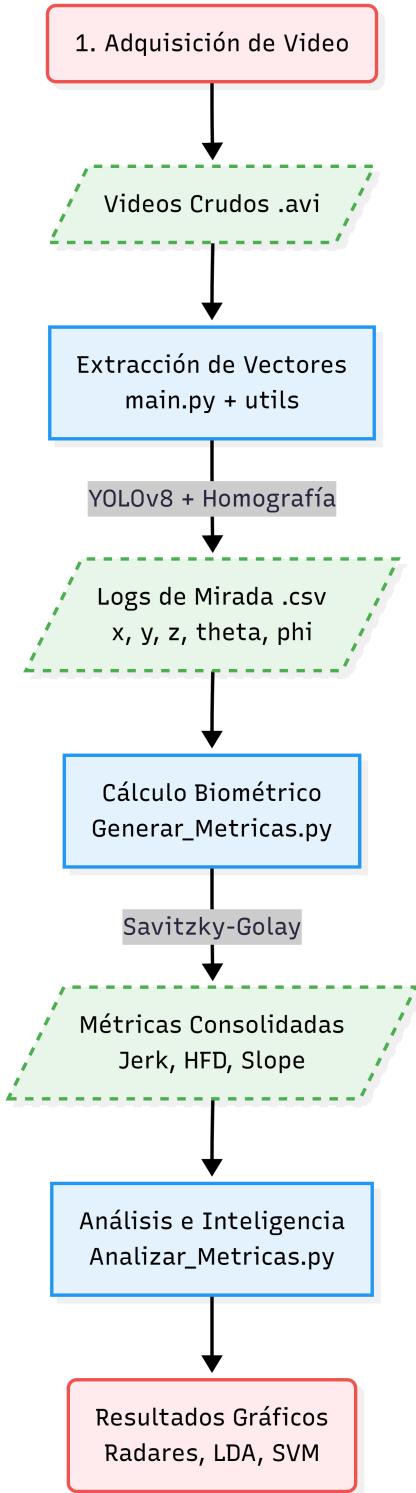


Figura 3.3: Diagrama de flujo de la arquitectura de procesamiento de datos. Se ilustra la secuencia desde la captura de video hasta la generación de resultados biométricos.

Adicionalmente, la Tabla 3.1 detalla la función específica de cada módulo de software dentro del repositorio del proyecto, especificando las entradas y salidas de cada etapa.

Tabla 3.1: Descripción funcional de los módulos de software del sistema.

| Etapa | Módulo (Script) | Entrada (Input) | Salida (Output) |
|----------------------|---------------------------------|--|--|
| 1. Extracción | main.py eye_tracker_utils.py | Videos del ojo (120 FPS) | Archivos _data.csv con centroides y vector mirada (g_x, g_y, g_z). |
| 2. Métricas | Generar_Metricas.py | Series temporales de posición ($P(t)$) | Dataset BIOMETRIC.csv con 18 descriptores (Jerk, HFD, etc.) |
| 3. Análisis | Analizar_Metricas.py | Dataset Biométrico consolidado | Gráficos de Radar, Clusters LDA y Matrices de Confusión. |

3.3. Protocolo de Adquisición de Datos

3.3.1. Población y Muestra

Caracterización de la población objetivo

La población objetivo de este estudio la conforman adultos jóvenes sin patologías oculomotoras diagnosticadas, residentes en Mérida (Venezuela). Este grupo demográfico fue seleccionado por presentar características oculomotoras estables y bien documentadas en la literatura científica, lo que facilita la comparación de los resultados obtenidos con estudios previos.

Criterios de selección

Para garantizar la homogeneidad de la muestra y la validez de las mediciones, se establecieron los siguientes criterios de inclusión y exclusión:

Criterios de inclusión:

- Edad comprendida entre 20 y 35 años
- Ausencia de patologías oculares diagnosticadas (cataratas, glaucoma, desprendimiento de retina)
- Comprensión de las instrucciones del experimento

Tamaño y composición de la muestra

La muestra final consistió en 15 participantes ($N = 15$), seleccionados mediante muestreo no probabilístico por conveniencia. La composición demográfica fue:

- **Edad:** Media de 25.3 años, desviación estándar de 3.7 años
- **Sexo:** 11 hombres (73.3%) y 4 mujeres (26.7%)
- **Corrección visual:** 8 participantes con visión normal (53.3%) y 7 con corrección óptica (lentes o lentes de contacto, 46.7%)

Este tamaño de muestra, aunque limitado, es consistente con estudios piloto en el campo de la biometría oculomotora. Investigaciones previas como las de Komogortsev et al. [1] han demostrado que muestras de 10-20 participantes son suficientes para validar la viabilidad de sistemas de seguimiento ocular y establecer la singularidad de patrones oculomotores individuales en condiciones controladas. A continuación, la Tabla 3.2 detalla las condiciones específicas registradas para cada uno de los 15 participantes durante las sesiones experimentales.

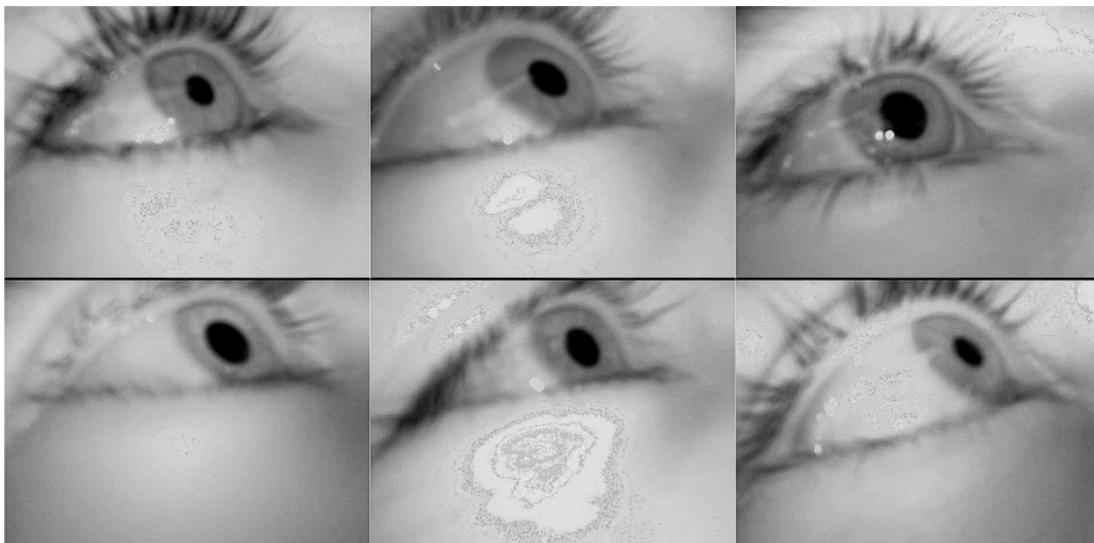


Figura 3.4: Muestra representativa de los ojos capturados durante la sesión experimental. Cada imagen corresponde a un participante distinto del estudio ($N = 15$), evidenciando la variabilidad morfológica inter-sujeto en características como el diámetro pupilar y rasgos que el sistema explota como descriptores biométricos primarios.

Tabla 3.2: Caracterización demográfica y condiciones ambientales por participante con incertidumbres asociadas.

| Sujeto | Edad (años) | Distancia (cm) | Iluminancia (Lux) |
|--------|-------------|----------------|-------------------|
| P01 | 24 | 59 | 585 |
| P02 | 26 | 60 | 620 |
| P03 | 22 | 60 | 550 |
| P04 | 29 | 58 | 710 |
| P05 | 25 | 60 | 680 |
| P06 | 31 | 60 | 850 |
| P07 | 23 | 59 | 595 |
| P08 | 27 | 60 | 640 |
| P09 | 21 | 59 | 520 |
| P10 | 35 | 60 | 890 |
| P11 | 28 | 60 | 730 |
| P12 | 24 | 60 | 615 |
| P13 | 30 | 59 | 800 |
| P14 | 22 | 60 | 560 |
| P15 | 26 | 60 | 675 |

Consideraciones Metrológicas:

- **Distancia (± 1 cm):** Medición realizada con cinta métrica comercial (resolución 1 mm). Se asigna una incertidumbre expandida de ± 1 cm para compensar el error de posicionamiento de la cabeza (paralaje y micro-movimientos).
- **Iluminancia ($\pm \Delta$):** Medición realizada con luxómetro digital UNI-T UT383. Según especificaciones del fabricante, la precisión instrumental en el rango < 9999 Lux es de $\pm(4\% \text{ lectura} + 8 \text{ dígitos})$.

Consideraciones éticas

Todos los participantes fueron informados sobre los objetivos del estudio, los procedimientos a seguir, y el uso que se daría a los datos recolectados. Se garantizó la confidencialidad de la información personal mediante la asignación de códigos anónimos (P01-P15). No se registraron imágenes que permitieran la identificación facial de los participantes, únicamente las regiones oculares necesarias para el análisis.

3.3.2. Diseño del Estímulo Visual

Cuadrícula de calibración y prueba (Grid 3x3)

El estímulo visual consistió en una cuadrícula regular de 3×3 puntos distribuidos uniformemente sobre el área activa del monitor. La posición de los puntos se calculó dividiendo la resolución de pantalla ($W \times H$) en tres segmentos iguales tanto horizontal como verticalmente, situando cada estímulo en el centro geométrico de cada celda resultante. Para una resolución estándar de 1920×1080 píxeles, esto resulta en una separación constante entre centros de 640 píxeles en el eje horizontal y 360 píxeles en el eje vertical.

Figura 3.5: Esquema de la cuadrícula de estímulos visuales 3×3 . Los números indican el orden secuencial de presentación. Para una resolución de 1920×1080 , la separación horizontal entre estímulos (Δx) es de 640 px y la vertical (Δy) es de 360 px.

Características geométricas y cromáticas del estímulo:

- **Número de puntos:** 9 (matriz de 3×3)
- **Forma y dimensión:** Círculos sólidos con un radio de 30 píxeles (diámetro total de 60 píxeles). A una distancia de observación aproximada de 60 cm, esto subtende un ángulo visual de $\approx 1.5^\circ$, garantizando una estimulación foveal clara.
- **Cromaticidad del estímulo:** Rojo puro de máxima intensidad (Espacio de color BGR: 0, 0, 255; RGB: 255, 0, 0).
- **Fondo:** Negro absoluto (RGB: 0, 0, 0). Se seleccionó un fondo de mínima luminancia para maximizar el contraste con el estímulo y eliminar distracciones visuales periféricas, facilitando la segmentación de la pupila al reducir reflejos externos en la córnea.
- **Distribución espacial:**
 - Separación Horizontal: $W/3$ (640 px), induciendo sacádicos de amplitud media-grande ($\approx 15^\circ - 20^\circ$).
 - Separación Vertical: $H/3$ (360 px), induciendo sacádicos verticales de amplitud controlada.

Esta configuración espacial fuerza al sistema oculomotor a realizar movimientos sacádicos de amplitudes conocidas, cubriendo tanto el rango lineal como la región de saturación de la *Main Sequence* (Ecuación 2.1), lo cual es fundamental para validar la dinámica del movimiento ocular capturado.

Secuencia temporal de presentación

Los estímulos se presentaron siguiendo una secuencia determinista controlada por software, garantizando la repetibilidad del experimento entre sujetos:

1. **Patrón de barrido:** La activación siguió un orden de lectura occidental (izquierda a derecha, arriba hacia abajo):

- Fila superior ($y = H/6$): Posiciones 1, 2, 3
- Fila media ($y = H/2$): Posiciones 4, 5, 6
- Fila inferior ($y = 5H/6$): Posiciones 7, 8, 9

Este patrón reduce la carga cognitiva del participante al hacer predecible la ubicación del siguiente objetivo, permitiendo enfocarse puramente en la tarea motora visual.

2. **Temporización estricta:** Cada estímulo permaneció estático en pantalla durante un intervalo exacto de $T = 2000$ ms (2 segundos). Este intervalo se seleccionó considerando la fisiología del ojo:

- *Latencia sacádica:* ≈ 200 ms.
- *Tiempo de vuelo:* 30 – 100 ms.
- *Fijación estable:* El tiempo restante (> 1.7 s) asegura la captura de suficientes fotogramas estables para el cálculo preciso del centroide pupilar promedio en cada posición.

3. **Sincronización de eventos:** El cambio de posición del estímulo se programó para ser instantáneo (transición en el siguiente refresco de pantalla), generando un estímulo tipo escalón (step stimulus) ideal para medir la respuesta al impulso del sistema oculomotor.

3.3.3. Procedimiento Experimental

El protocolo experimental se diseñó para garantizar condiciones estandarizadas y replicables entre todos los participantes ($N = 15$). Cada sesión individual tuvo una duración aproximada de 10 minutos e incluyó las siguientes etapas secuenciales:

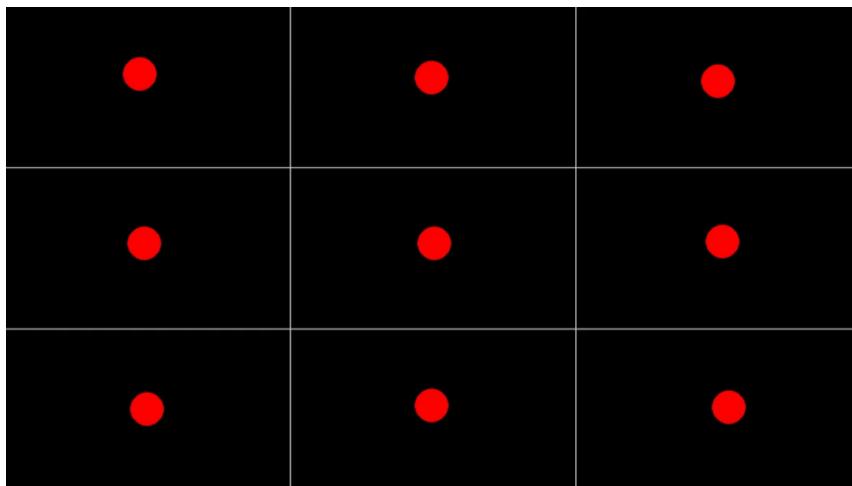


Figura 3.6: Visualización del protocolo de estímulos (Experimento 1). La matriz de 3×3 puntos rojos sobre fondo negro se utiliza para inducir movimientos sacádicos controlados. Los puntos aparecen de forma secuencial con una duración de 2 segundos por posición, cubriendo la totalidad del campo de visión efectivo.

Fase 1: Recepción y consentimiento informado

1. **Bienvenida:** Se recibió al participante y se le explicó verbalmente el propósito general del estudio: analizar la dinámica oculomotora mediante visión artificial.
2. **Registro de metadatos:** Se completó una ficha técnica con las variables de control: Edad, Sexo y Tipo de corrección visual (gafas/lentes de contacto), dado que el 46.7% de la muestra utilizaba algún tipo de ayuda óptica.

Fase 2: Colocación del dispositivo y configuración óptica

Dado que el sistema de captura es del tipo *head-mounted* (montado en la cabeza), el procedimiento de ajuste difiere de los sistemas remotos tradicionales:

1. **Colocación del dispositivo:** El participante se colocó la montura de gafas que soporta la cámara GC0308. Se verificó que la estructura fuera cómoda y estable, asegurando que la cámara quedara ubicada en el cuadrante inferior derecho del campo visual, apuntando hacia el ojo en un ángulo de elevación (aprox. 30°).
2. **Ajuste del ROI (Región de Interés):** Mediante la visualización en tiempo real en el monitor de control, se ajustó mecánicamente el brazo flexible de la cámara para centrar el ojo en la imagen de 320×240 píxeles, garantizando que la pupila se mantuviera dentro del encuadre incluso durante movimientos extremos hacia las esquinas.

3. **Posicionamiento frente al estímulo:** El participante se sentó frente al monitor. Se ajustó la altura de la silla para alinear los ojos con el centro de la pantalla y se fijó la distancia de observación a 60 cm (medida con cinta métrica) para asegurar que la geometría de los movimientos sacádicos correspondiera a los ángulos visuales calculados ($\approx 20^\circ$ horizontal).

Fase 3: Ejecución de la prueba principal

El software de control (desarrollado en Python) gestionó la sincronización entre el estímulo y la captura:

1. **Inicialización y Logs:** Al iniciar el script, se generó automáticamente un archivo CSV para registrar las marcas temporales (*timesteps*) de cada cambio de estímulo.
2. **Cuenta regresiva:** Se presentó una cuenta visual en pantalla (3-2-1) para preparar al sujeto.
3. **Protocolo de adquisición:** Se ejecutó la matriz de 9 puntos. El sistema grabó video continuo a 120 FPS en formato MJPEG para evitar latencia, mientras registraba simultáneamente en el log la posición (x, y) del punto rojo y el tiempo exacto de aparición.
4. **Control de parpadeo:** Se instruyó a los participantes intentar no parpadear durante los 2 segundos de fijación activa, permitiéndolo libremente durante las transiciones si fuera necesario.

Fase 4: Repetición y consistencia de datos

Para garantizar la robustez estadística y filtrar posibles artefactos (parpadeos involuntarios o pérdida de tracking), se realizó un diseño iterativo:

1. Se realizaron 3 iteraciones (intentos) completas del experimento por cada participante.
2. Entre iteraciones se estableció un intervalo de descanso aproximado de 10 segundos.
3. El volumen total de datos adquiridos se calcula como:

$$\text{Frames totales} = N_{\text{sujetos}} \times N_{\text{intentos}} \times (T_{\text{ensayo}} \times FPS)$$

$$15 \times 3 \times (18 \text{ s} \times 120 \text{ fps}) \approx 97,200 \text{ imágenes oculares}$$

Esta alta densidad temporal (120 Hz) proporciona suficiente información para reconstruir la curva de velocidad de los movimientos sacádicos con gran detalle.

Fase 5: Cierre de la sesión

1. Se retiró el dispositivo del participante.
2. Se verificó la integridad de los archivos de video (.avi/.mp4) y los logs de datos (.csv) antes de liberar al sujeto.
3. Se agradeció la participación y se procedió a preparar el equipo de captura para el siguiente usuario.

3.4. Procesamiento Digital de Imágenes

El procesamiento digital de imágenes constituye el núcleo del sistema de seguimiento ocular desarrollado. Esta sección describe, desde una perspectiva algorítmica y matemática, cómo se transforman las imágenes RGB capturadas por la cámara en coordenadas precisas del centroide de la pupila. El pipeline de procesamiento se divide en tres etapas fundamentales: preprocesamiento, segmentación y detección de la pupila, y mapeo de coordenadas.

3.4.1. Preprocesamiento

El objetivo del preprocesamiento es mejorar la calidad de las imágenes capturadas, reducir el ruido, y preparar los datos para los algoritmos de segmentación subsecuentes. Las operaciones de preprocesamiento se aplican secuencialmente a cada fotograma capturado.

Conversión a escala de grises

Las imágenes RGB capturadas por la cámara son tensores tridimensionales de dimensión $H \times W \times 3$, donde $H = 720$ píxeles (altura), $W = 1280$ píxeles (ancho), y los tres canales corresponden a Rojo (R), Verde (G), y Azul (B). Para simplificar el procesamiento posterior y reducir la carga computacional, las imágenes se convierten a escala de grises mediante la transformación:

$$I_{\text{gray}}(x, y) = 0.2989 \cdot R(x, y) + 0.5870 \cdot G(x, y) + 0.1140 \cdot B(x, y) \quad (3.1)$$

donde $I_{\text{gray}}(x, y)$ es la intensidad en escala de grises del píxel ubicado en la posición (x, y) , y los coeficientes reflejan la sensibilidad espectral del sistema visual humano (el canal verde contribuye más a la percepción de luminancia). Esta transformación reduce el espacio de representación de 3 canales a 1 canal, pasando de $720 \times 1280 \times 3 = 2,764,800$ valores por fotograma a $720 \times 1280 =$

921,600 valores, lo cual acelera significativamente las operaciones subsecuentes sin pérdida de información relevante para la detección de la pupila.

Reducción de ruido mediante filtrado

Las imágenes capturadas por cámaras digitales inevitablemente contienen ruido, proveniente de fuentes como la sensibilidad del sensor, condiciones de baja iluminación, y artefactos de compresión. Para suavizar estas fluctuaciones aleatorias de intensidad sin perder información estructural importante (como los bordes de la pupila), se aplicaron dos tipos de filtros complementarios:

Filtro Gaussiano El filtro Gaussiano es un filtro lineal de paso bajo que atenúa las componentes de alta frecuencia espacial (ruido) mediante una convolución con un kernel gaussiano bidimensional descrito por Giménez et al. [13]:

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (3.2)$$

donde σ es la desviación estándar de la distribución gaussiana, que controla el grado de suavizado. En este estudio se utilizó $\sigma = 1.5$ y un kernel de tamaño 5×5 píxeles. La imagen filtrada se obtiene mediante:

$$I_{\text{filtrada}}(x,y) = (I_{\text{gray}} * G)(x,y) = \sum_{i=-2}^2 \sum_{j=-2}^2 I_{\text{gray}}(x-i, y-j) \cdot G(i,j) \quad (3.3)$$

3.4.2. Segmentación y Detección de Pupila

El proceso de segmentación se implementó mediante un flujo de trabajo (pipeline) de visión artificial diseñado para operar con alta eficiencia temporal (120 FPS). A diferencia de los métodos clásicos basados puramente en procesamiento de imagen, este estudio integra un modelo de aprendizaje profundo para la localización robusta, seguido de algoritmos geométricos para la precisión sub-pixel.

Localización y Seguimiento de la Pupila mediante Deep Learning

Para la localización robusta de la pupila frame a frame, se implementó una red neuronal convolucional basada en la arquitectura YOLOv8 (*You Only Look Once*, versión 8). A diferencia de los enfoques genéricos de detección de rostros, se optó por entrenar un modelo específico para la detección de la pupila en entornos infrarrojos cercanos (NIR).

Arquitectura y Entrenamiento (Fine-Tuning) Se seleccionó la variante YOLOv8n-seg (Nano), la arquitectura más ligera de la familia YOLO (3.2 millones de parámetros), para garantizar una inferencia en tiempo real compatible con la tasa de captura de 120 FPS. El modelo no se utilizó de caja (*out-of-the-box*), sino que se sometió a un proceso de transferencia de aprendizaje (*fine-tuning*):

1. **Conjunto de datos (Dataset):** Se construyó un dataset personalizado utilizando fotografías extraídos de los propios participantes del estudio durante la fase de calibración. Se etiquetaron manualmente 864 imágenes representativas que incluían variaciones en la apertura palpebral, reflejos corneales y distintas iluminaciones.
2. **Configuración del entrenamiento:** El modelo pre-entrenado en el dataset COCO se re-entrenó durante 80 épocas con un tamaño de lote (batch size) de 16 y un optimizador SGD (Descenso de Gradiente Estocástico) con momentum de 0.937.
3. **Validación:** El modelo resultante (`best.pt`) alcanzó una precisión media (mAP@50) superior al 98 % en el conjunto de validación, demostrando una capacidad de generalización robusta frente a la variabilidad inter-sujeto.

Esta estrategia de entrenamiento específico resultó fundamental para mitigar la pérdida de seguimiento durante parpadeos o movimientos sacádicos rápidos, superando las limitaciones de los métodos tradicionales de visión artificial.

Inferencia en tiempo de ejecución Durante la fase experimental, el algoritmo opera de la siguiente manera para cada fotograma I_t :

1. **Pre-recorte:** Se ajusta la relación de aspecto de la imagen de entrada a 320×240 para mantener la consistencia espacial con los datos de entrenamiento.
2. **Inferencia:** El modelo predice un cuadro delimitador (bounding box) $B = (x_1, y_1, x_2, y_2)$ centrado en la pupila.
3. **Filtrado y Expansión:** Se aceptan únicamente detecciones con una confianza $C \geq 0.7$. Las coordenadas predichas se expanden en 5 píxeles por lado (*padding*) para definir la Región de Interés (ROI) final, asegurando que los bordes de la pupila se preserven íntegramente para la etapa de segmentación sub-pixel.

Pre-procesamiento y mejora de contraste

Dado que la iluminación infrarroja puede generar histogramas de intensidad concentrados en rangos oscuros, se aplicaron técnicas de mejora de imagen antes de la segmentación. Según se observa en la función `process_frame` del algoritmo implementado:

1. **Suavizado Gaussiano:** Se aplicó un filtro Gaussiano con kernel (7×7) para reducir el ruido de alta frecuencia del sensor CMOS.
2. **Ecuación de Histograma Adaptativa (CLAHE):** Se utilizó *Contrast Limited Adaptive Histogram Equalization* con un límite de recorte (*clip limit*) de 1.0 y una rejilla de 8×8 . Esto maximiza el contraste local entre la pupila y el iris sin amplificar el ruido en las zonas homogéneas, lo cual es crítico para el método de pupila oscura.

Segmentación y Ajuste Geométrico

En lugar de depender de operaciones morfológicas tradicionales que pueden deformar la forma de la pupila, se implementó un enfoque basado en la geometría de contornos:

1. **Binarización** Se aplicó una umbralización binaria invertida fija sobre la imagen mejorada por CLAHE:

$$I_{\text{bin}}(x, y) = \begin{cases} 255 & \text{si } I_{\text{CLAHE}}(x, y) < T_{\text{fijo}} \\ 0 & \text{en caso contrario} \end{cases} \quad (3.4)$$

Donde T_{fijo} es un parámetro empírico (configurado en 14-20 niveles de intensidad) que separa la región negra profunda de la pupila del resto del ojo.

2. **Optimización de Contornos por Ángulo** Tras detectar los contornos, se seleccionó el contorno de mayor área. Para eliminar irregularidades causadas por pestañas o reflejos (glints), se implementó un algoritmo de filtrado geométrico (`optimize_contours_by_angle`). Este algoritmo evalúa la suavidad de la curvatura calculando el coseno del ángulo entre vectores formados por puntos adyacentes del contorno:

$$\cos \theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|} \quad (3.5)$$

Los puntos que generan cambios angulares abruptos (vértices agudos no compatibles con la curvatura de la pupila) son descartados como ruido.

Algoritmo 1 Optimización de Contornos por Suavidad Angular

Entrada: Contorno original $C = \{P_1, P_2, \dots, P_N\}$, Umbral de suavidad τ

Salida: Contorno optimizado C_{opt}

```
1:  $C_{opt} \leftarrow \emptyset$ 
2: for  $i \leftarrow 1$  to  $N$  do
3:   Obtener puntos adyacentes:
4:    $P_{prev} \leftarrow C[(i - 1) \text{ (mód } N)]$ 
5:    $P_{curr} \leftarrow C[i]$ 
6:    $P_{next} \leftarrow C[(i + 1) \text{ (mód } N)]$ 
7:   Calcular vectores direccionales:
8:    $\vec{v}_1 \leftarrow P_{curr} - P_{prev}$ 
9:    $\vec{v}_2 \leftarrow P_{next} - P_{curr}$ 
10:  Calcular similitud angular (Coseno):
11:   $\cos \theta \leftarrow \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|}$ 
12:  Filtrado geométrico:
13:  if  $\cos \theta > \tau$  then                                ▷ El ángulo es suave (compatible con curvatura)
14:    Agregar  $P_{curr}$  a  $C_{opt}$ 
15:  else
16:    Descartar  $P_{curr}$                                 ▷ Vértice agudo (ruido/pestaña)
17:  end if
18: end for
19: return  $C_{opt}$ 
```

3. Ajuste de Elipse (Least Squares Fitting) Finalmente, se ajustó una elipse a los puntos del contorno optimizado utilizando el método de mínimos cuadrados directos (*Direct Least Squares fitting of Ellipses*). Esto retorna el centroide con precisión sub-pixel (x_c, y_c), los ejes mayor y menor, y el ángulo de rotación, proporcionando una estimación robusta de la posición de la pupila incluso ante occlusiones parciales.

3.4.3. Estimación del Vector de Mirada (Gaze Vector)

A diferencia de los mapeos polinomiales 2D simples, este estudio implementó un modelo geométrico 3D simplificado para estimar la dirección de la mirada. El algoritmo calcula un Centro del Modelo (aproximación del centro de rotación ocular o centro óptico virtual) mediante la intersección promedio de los vectores normales a la superficie del ojo a lo largo del tiempo. Posteriormente, tal como se ilustra en la Figura 3.8, el vector de mirada unitario \hat{g} se calcula

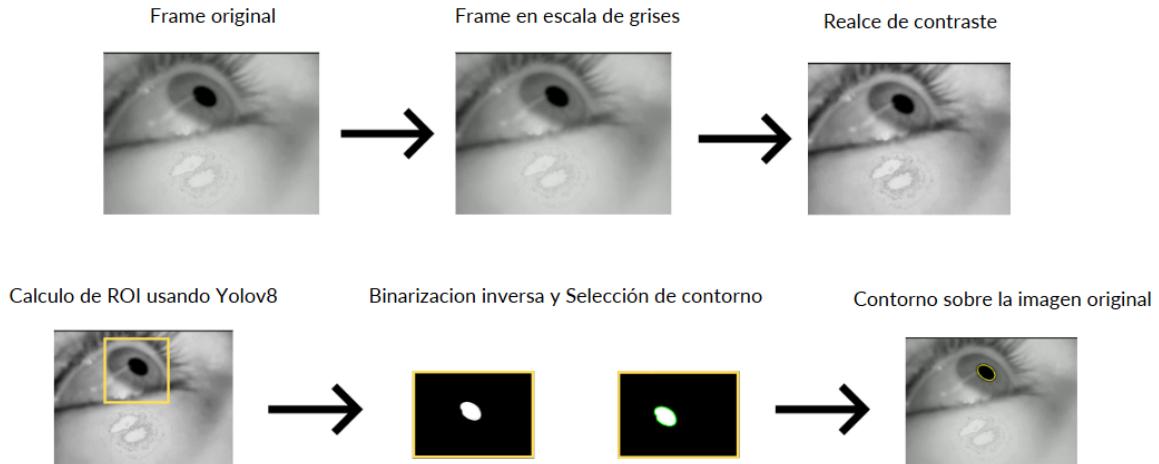


Figura 3.7: Pipeline de procesamiento de imagen implementado. La secuencia muestra: (Arriba) Mejora de la imagen mediante escala de grises y realce de contraste CLAHE. (Abajo) Detección de la ROI con YOLOv8, binarización inversa adaptativa y superposición final del contorno elíptico ajustado sobre la imagen original.

normalizando la diferencia entre el centro de la pupila detectado $\mathbf{p} = (x_p, y_p, 0)$ y el centro del modelo $\mathbf{c} = (x_c, y_c, z_c)$:

$$\vec{v} = \mathbf{p} - \mathbf{c}, \quad \hat{g} = \frac{\vec{v}}{\|\vec{v}\|} = (g_x, g_y, g_z) \quad (3.6)$$

Este enfoque vectorial permite calcular la cinemática angular de manera más fiel a la fisiología del ojo.

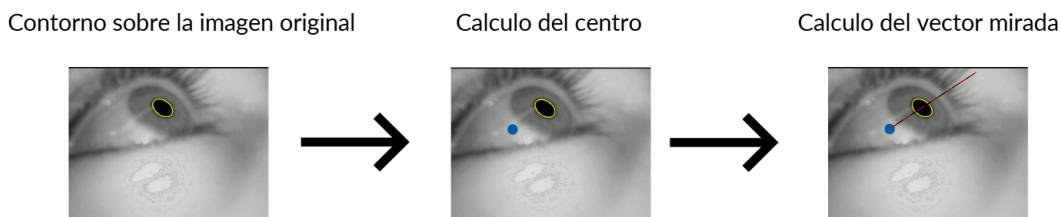


Figura 3.8: Esquema de la estimación del vector de mirada. (Izquierda) Identificación del contorno pupilar. (Centro) Cálculo del centro del modelo ocular (punto azul) respecto al centroide de la pupila. (Derecha) Proyección del vector de mirada resultante (línea roja) en el espacio 3D.

Validación y Justificación del Enfoque Híbrido

La selección de una arquitectura basada en aprendizaje profundo (YOLOv8) frente a los métodos clásicos de procesamiento de imagen se fundamenta en pruebas preliminares de rendimiento y robustez realizadas durante el diseño del sistema.

Robustez en la detección En implementaciones previas basadas puramente en técnicas de visión artificial (operaciones morfológicas y umbralización global sobre el fotograma completo), se observó una tasa de fallo en la detección de la pupila de aproximadamente un 10% de los fotogramas procesados. Estas pérdidas ocurrían principalmente debido a:

- Oclusiones parciales por pestañas o párpados durante el parpadeo.
- Cambios bruscos de iluminación que afectaban el umbral de binarización.
- Pérdida de seguimiento durante movimientos sacádicos de alta velocidad.

Por el contrario, la implementación propuesta mediante YOLOv8 demostró una estabilidad superior, con menos del 1% de fallos en la detección de la ROI en condiciones controladas. El modelo es capaz de re-adquirir la posición del ojo instantáneamente frame a frame específicamente para el dataset de esta investigación, garantizando la continuidad de las series temporales necesarias para el cálculo de la velocidad y aceleración.

Eficiencia Computacional Además de la robustez, el uso de YOLO para recortar la Región de Interés (ROI) ofrece una ventaja crítica de optimización. Al limitar el procesamiento matemático costoso (ajuste de elipses, CLAHE, algoritmos de contornos) exclusivamente al área delimitada por el *bounding box* (< 10% del área total de la imagen), se reduce drásticamente la carga computacional. Esto libera recursos del CPU para mantener una tasa de muestreo estable de 120 FPS, algo que sería insostenible procesando la matriz de imagen completa con algoritmos iterativos.

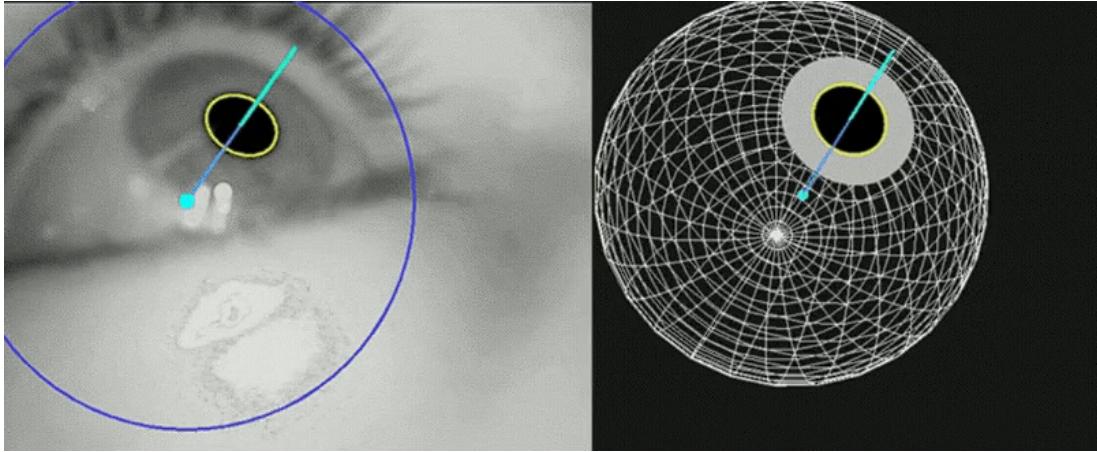


Figura 3.9: Correspondencia entre la señal visual capturada y su representación computacional.
Izquierda: imagen infrarroja del ojo adquirida por la cámara, donde se aprecia la pupila detectada.
Derecha: modelo tridimensional resultante, en el que el diámetro pupilar estimado determina la escala del globo ocular y el vector de mirada (*gaze vector*) que define su orientación en el espacio 3D. Esta transformación constituye el núcleo del pipeline de estimación oculomotora propuesto.

3.5. Procesamiento de Señales y Extracción de Métricas

El archivo `Generar_Metricas.py` procesó las series temporales de vectores de mirada $\hat{g}(t)$ para extraer los descriptores biométricos.

3.5.1. Cálculo de la Cinemática Angular

Para obtener la posición angular $\theta(t)$ (en grados) a partir de los vectores unitarios de mirada, se utilizó el producto punto entre vectores consecutivos, lo cual mide el desplazamiento angular geodésico independiente de la geometría de la pantalla:

$$\Delta\theta_i = \arccos(\hat{g}_i \cdot \hat{g}_{i-1}) \cdot \frac{180}{\pi} \quad (3.7)$$

La trayectoria angular acumulada se define como $P(t) = \sum \Delta\theta$.

3.5.2. Filtrado y Derivación

Para el cálculo de la velocidad y aceleración, se aplicó el filtro Savitzky-Golay sobre la señal de posición angular $P(t)$, con los siguientes parámetros configurados en el sistema (CONFIG ['PARAMS']):

- **Ventana (w):** 21 muestras (equivalente a 175 ms a 120 Hz).
- **Orden del polinomio (p):** 3 (cúbico).

Estos parámetros se seleccionaron mediante optimización empírica, evaluando diferentes combinaciones hasta encontrar el mejor balance entre suavizado y preservación de características. Las derivadas se obtuvieron analíticamente a partir de los coeficientes del polinomio ajustado, lo que reduce significativamente la amplificación de ruido comparado con las diferencias finitas:

$$V(t) = \frac{d}{dt} \text{SavGol}(P(t)) \quad [\text{°}/\text{s}] \quad (3.8)$$

$$a(t) = \frac{d}{dt} V(t) \quad [\text{°}/\text{s}^2] \quad (3.9)$$

3.5.3. Nuevas Métricas Biométricas Integradas

Además de las métricas estándar, se implementaron algoritmos para extraer características avanzadas de control motor y cognitivo:

Pendiente de la Secuencia Principal (Main Sequence Slope)

Para caracterizar la biomecánica muscular, se analizó la relación entre la velocidad pico (V_{pico}) y la amplitud (Amp) de las sacadas. En el rango de amplitudes medidas, esta relación se linealizó, calculando la pendiente K mediante regresión lineal:

$$V_{pico} \approx K \cdot Amp \implies K = \frac{\text{Cov}(V_{pico}, Amp)}{\text{Var}(Amp)} \quad (3.10)$$

Este parámetro K es un indicador de la rigidez o eficiencia del sistema oculomotor.

Dimensión Fractal de Higuchi (HFD)

Para cuantificar la complejidad de la señal de velocidad (estrategia cognitiva de búsqueda), se aplicó el algoritmo de Higuchi con un parámetro $k_{max} = 5$. Este algoritmo calcula la dimensión fractal D_H basándose en la tasa de cambio de la longitud de la curva $L(k)$ a diferentes escalas temporales k :

$$L(k) \propto k^{-D_H} \quad (3.11)$$

Un valor de D_H más alto indica una señal más compleja y caótica, mientras que un valor bajo indica movimientos más deterministas y suaves.

Velocidad Pupilar Dinámica

Se calculó la derivada temporal del diámetro pupilar para obtener la velocidad de contracción/dilatación. La métrica `Pupil_Vel_Max` captura la reactividad máxima del sistema nervioso autónomo ante la carga cognitiva del estímulo.

3.5.4. Justificación y Selección del Espacio de Características

La selección de los descriptores biométricos no fue arbitraria: cada métrica fue elegida por su capacidad para capturar una dimensión fisiológica independiente del comportamiento oculomotor, maximizando la información biométrica disponible mientras se minimiza la redundancia entre variables. A continuación se justifica teóricamente el conjunto de 15 descriptores organizados en cuatro categorías funcionales, y se explica el criterio de selección automática posterior mediante el índice de impureza de Gini.

Categoría I — Cinemática Sacádica

Las métricas de esta categoría describen la dinámica del movimiento sacádico a distintas escalas de derivación temporal:

- `Vel_Mean`
- `Acc_Max`
- `Jerk_Mean`
- `Jerk_Max`
- `Velocity_Transition_Smoothness`

Su justificación biométrica se fundamenta en el trabajo de Komogortsev et al. [1], quienes demostraron que los parámetros del modelo de la planta oculomotora —incluyendo la velocidad y la aceleración— son lo suficientemente estables inter-sesión e idiosincrásicos inter-sujeto como para constituir una firma biométrica. En particular, el *Jerk* (derivada de la aceleración, $J = \ddot{\theta}$) es especialmente relevante porque cuantifica el costo energético del movimiento:

$$J(t) = \frac{d^3\theta}{dt^3} \quad [\text{°}/\text{s}^3] \quad (3.12)$$

El cerebro minimiza el *jerk* cuadrático integrado al planificar movimientos oculares óptimos [20], y la forma exacta en que cada individuo alcanza ese mínimo está determinada por la geometría muscular única de su sistema oculomotor [2]. Así, `Jerk_Max` captura el límite

superior del esfuerzo neuromuscular, mientras que `Jerk_Mean` describe la estrategia promedio de control motor. La métrica `Velocity_Transition_Smoothness` complementa estas variables midiendo la continuidad de la curva de velocidad, un indicador de la calidad del control motor de lazo cerrado que varía entre individuos con distinta coordinación oculomotora [21].

Categoría II — Morfología Pupilar

Las métricas `Pupil_Mean`, `Pupil_Std`, `Pupil_CV` y `Pupil_Vel_Max` caracterizan la geometría y la dinámica del diámetro pupilar. La justificación de su inclusión es doble.

En primer lugar, el diámetro pupilar en reposo es un rasgo fenotípico estable determinado por la anatomía del músculo esfínter del iris [16]. Su valor medio (`Pupil_Mean`) actúa como un descriptor morfológico, análogo al diámetro del iris en los sistemas de reconocimiento biométrico clásico [1]. El coeficiente de variación ($Pupil_{CV} = \sigma/\mu$) normaliza la dispersión por la media, haciéndolo comparable entre sujetos con distintos tamaños pupilares basales:

$$CV_{pupil} = \frac{\sigma_{pupil}}{\mu_{pupil}} \times 100\% \quad (3.13)$$

En segundo lugar, la dinámica pupilar (`Pupil_Vel_Max`) refleja la reactividad del sistema nervioso autónomo ante la carga cognitiva del estímulo [17]. Beatty demostró que la dilatación pupilar evocada por la tarea (*task-evoked pupillary response*) es proporcional al esfuerzo cognitivo y presenta alta consistencia intra-sujeto, lo que la convierte en un marcador conductual de la estrategia de procesamiento visual individual.

Categoría III — Estructura Temporal del Movimiento

Las métricas de esta categoría describen la organización temporal de la exploración visual: cuántos movimientos se producen, con qué frecuencia, y qué tan eficientemente se realizan:

- `Microsaccade_Rate`
- `Saccade_Rate`
- `Fixation_Vel_Mean`
- `Main_Seq_Slope`

La tasa de microsacadas (`Microsaccade_Rate`) es particularmente relevante como descriptor de biometría conductual: los microsacadas son movimientos involuntarios de fijación que ocurren durante la fixación ocular, y su frecuencia y dirección están moduladas por factores atencionales y neurofisiológicos propios de cada individuo [22]. La pendiente de la Secuencia

Principal (Main_Seq_Slope) cuantifica la eficiencia neuromuscular del sistema sacádico como fue descrito en la Sección 3.5.4, y su inclusión como métrica independiente permite capturar variaciones individuales en la curva velocidad-amplitud que el modelo exponencial completo de Bahill [2] no resume en un único escalar. La velocidad durante fijación (Fixation_Vel_Mean) captura la micro-deriva residual del ojo durante los períodos de aparente reposo, un marcador de la estabilidad del sistema de mantenimiento de fijación que varía entre individuos sanos [20].

Categoría IV — Complejidad No Lineal

Las métricas Fractal_Dim y Gaze_Z_Mean aportan información de orden superior que los descriptores lineales no pueden capturar.

La Dimensión Fractal de Higuchi (Fractal_Dim) cuantifica la auto-similaridad de la señal de velocidad a través de múltiples escalas temporales [18]. Señales de mayor dimensión fractal (D_H próxima a 2) reflejan estrategias de exploración visual más complejas y caóticas, mientras que valores bajos (D_H próxima a 1) corresponden a movimientos más suaves y predecibles. Esta métrica es especialmente valiosa porque es invarianta a transformaciones lineales de la señal, lo que la hace robusta ante variaciones de ganancia del sensor. La componente Z del vector de mirada (Gaze_Z_Mean) codifica la profundidad de la dirección de mirada en el espacio tridimensional, capturando diferencias individuales en la postura de convergencia ocular que son invisibles para los sistemas de eye-tracking 2D convencionales [16].

Reducción del Espacio de Características mediante Impureza de Gini

El conjunto inicial de 15 descriptores fue sometido a un proceso de selección automática usando el índice de impureza de Gini calculado durante el entrenamiento del clasificador *Random Forest* [3]. El criterio de Gini mide, para cada característica X_j , la reducción promedio de impureza que produce en los nodos de decisión del ensemble:

$$\text{Importancia}(X_j) = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \sum_{n \in t} \mathbb{1}[X_j \text{ divide } n] \cdot \Delta\text{Gini}(n) \quad (3.14)$$

donde \mathcal{T} es el conjunto de árboles del ensemble y $\Delta\text{Gini}(n)$ es la reducción de impureza en el nodo n . Las características con importancia de Gini cercana a cero indican que, una vez que las demás variables han realizado sus particiones, dicha característica no aporta información adicional para separar las clases. Su eliminación no degrada el rendimiento del clasificador pero sí reduce el riesgo de sobreajuste (*overfitting*) al disminuir la dimensionalidad efectiva del espacio de decisión [3].

Este mecanismo de selección tiene dos ventajas metodológicas frente a métodos externos como la correlación de Pearson o el análisis de componentes principales (PCA): opera directa-

mente sobre el criterio de clasificación (no sobre la varianza de los datos), y es no paramétrico, por lo que detecta relaciones no lineales entre características y clases que los métodos lineales pasarían por alto. El resultado —reportado en la Figura 4.5 del Capítulo 4— constituye así tanto una herramienta de interpretación fisiológica como una regularización implícita del modelo de clasificación.

3.6. Resumen del Capítulo

En este capítulo se ha detallado la metodología del sistema, fundamentada en un enfoque híbrido que combina la robustez del aprendizaje profundo (YOLOv8) con la precisión de la óptica geométrica y el análisis de señales. Se ha descrito el diseño experimental controlado con 15 participantes, la configuración del hardware de captura de alta velocidad (120 FPS) y el pipeline de procesamiento que transforma imágenes crudas en descriptores biométricos complejos como el Jerk y la Dimensión Fractal. Esta metodología garantiza que los datos obtenidos sean precisos espacialmente y preserven la riqueza dinámica necesaria para caracterizar el comportamiento oculomotor. En el Capítulo 4 se presentarán los resultados obtenidos tras aplicar este procesamiento a la muestra recolectada, evaluando la capacidad de estas métricas para diferenciar patrones individuales y validar el rendimiento del sistema propuesto.

Capítulo 4

Resultados y Discusión

En este capítulo se presentan los resultados obtenidos tras el procesamiento y análisis de las señales oculomotoras de los 15 participantes del estudio. La exposición se organiza en cinco etapas fundamentales: primero, se valida la calidad técnica de la señal capturada y el rendimiento del algoritmo de detección; segundo, se caracteriza la dinámica fisiológica de los movimientos registrados; tercero, se evalúa la capacidad discriminativa de las métricas biométricas propuestas; cuarto, se presenta el rendimiento de los modelos de clasificación automática; y finalmente, se analiza la viabilidad del sistema para aplicaciones de control de cursor.

4.1. Validación del Sistema de Captura y Procesamiento

Antes de abordar el análisis biométrico, es fundamental verificar la integridad de los datos adquiridos y la robustez del sistema de visión artificial implementado.

4.1.1. Calidad de la Señal y Filtrado

El análisis inicial de los datos crudos reveló la presencia de ruido de alta frecuencia, un problema común en sensores CMOS cuando operan con ganancia variable en el espectro infrarrojo. Holmqvist et al. [22] señalan que este tipo de *jitter* instrumental es inherente a los sistemas de captura basados en cámaras no especializadas, y que su mitigación mediante filtrado digital es un paso indispensable antes de cualquier análisis cinematográfico, dado que las derivadas temporales (velocidad, aceleración) amplifican exponencialmente el ruido de alta frecuencia si se calculan sobre la señal cruda.

Para mitigar este *jitter* sin comprometer la integridad de la información biológica, se aplicó el filtro digital Savitzky-Golay. Savitzky y Golay [23] demostraron que el ajuste polinomial local sobre ventanas deslizantes preserva los momentos estadísticos de la señal —media, varianza y curtosis— con mayor fidelidad que los filtros de promedio móvil convencionales, precisamente porque el polinomio se adapta a la curvatura local de la señal en lugar de promediarla hacia cero. La configuración óptima del filtro se estableció con una ventana de longitud $w = 21$ muestras y un polinomio de orden $p = 3$. Esta elección de parámetros es crítica para el estudio:

- **Ventana ($w = 21$):** A una tasa de 120 FPS, esta ventana abarca un contexto temporal de ≈ 175 ms. Según el criterio de Nyquist [24], la frecuencia de muestreo de 120 Hz permite reconstruir fielmente componentes de señal de hasta 60 Hz, rango que cubre toda la dinámica fisiológica relevante del sistema oculomotor. La ventana de 175 ms proporciona suavizado suficiente para eliminar fluctuaciones aleatorias del centroide pupilar sin suprimir los eventos sacádicos, cuya duración mínima supera los 20 ms [2].
- **Polinomio Cúbico ($p = 3$):** A diferencia de los filtros de promedio móvil que tienden a atenuar o recortar los picos de señal, el ajuste polinomial de tercer grado preserva los momentos de inercia y, crucialmente, la magnitud real de la velocidad durante los movimientos sacádicos rápidos. Harwood, Mezey y Harris [25] confirmaron que el contenido frecuencial de los sacádicos se concentra por debajo de los 50 Hz, lo que hace al polinomio cúbico una aproximación suficientemente flexible para capturar esa dinámica sin sobreajustar el ruido.

Como se evidencia en la Figura 4.1, el filtro actúa de manera conservadora: elimina el ruido “sucio” de la señal cruda (línea negra) pero se adhiere perfectamente a las transiciones rápidas del ojo (línea de color), asegurando que no se eliminan micro-movimientos importantes ni se introduzca latencia artificial en la señal.

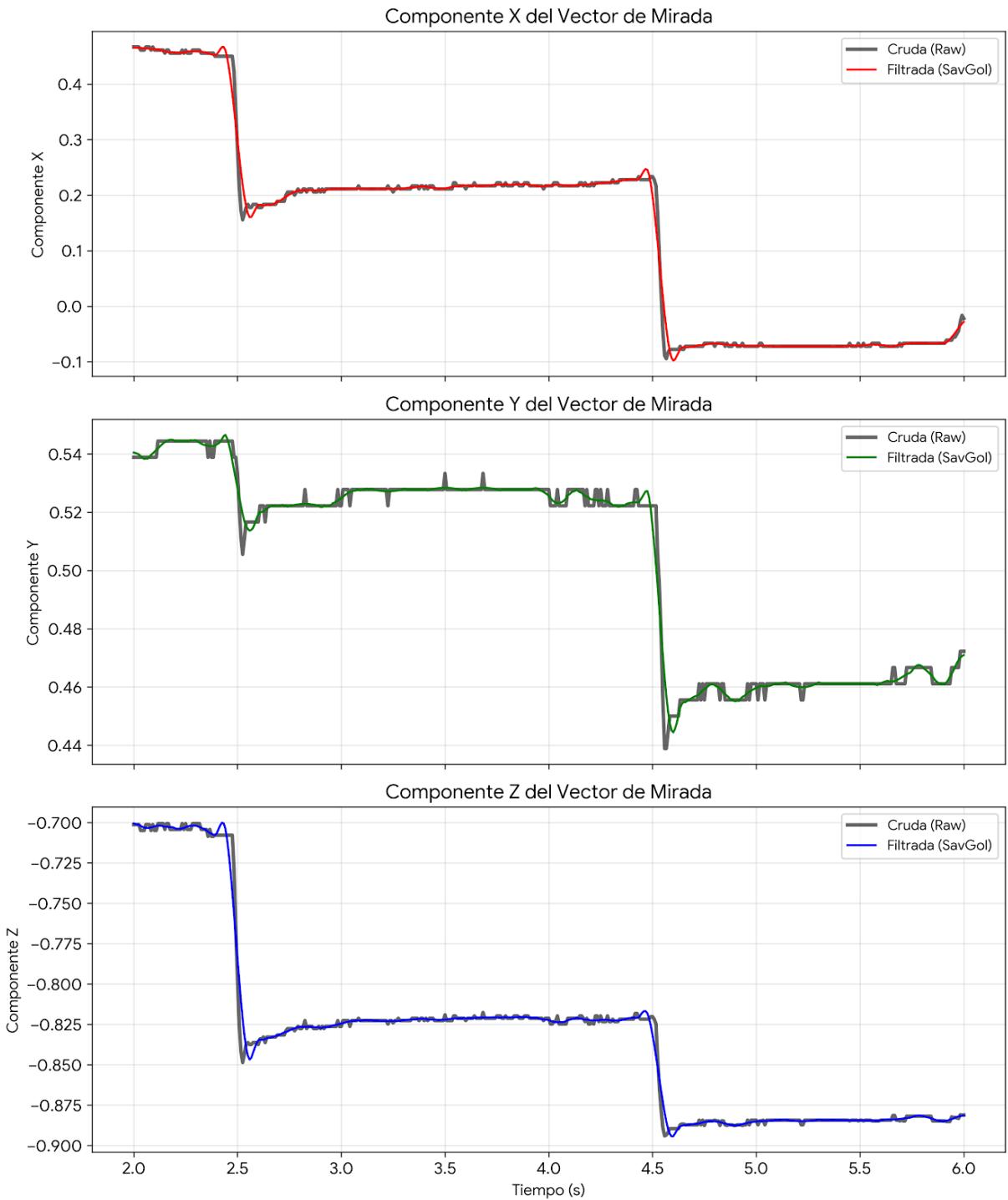


Figura 4.1: Descomposición vectorial de la señal de mirada en un corto periodo de tiempo. Se compara la señal cruda (negro) con la señal filtrada (colores) para las componentes X, Y y Z del vector de mirada. El filtro Savitzky-Golay ($w = 21$, $p = 3$) elimina el ruido de alta frecuencia manteniendo la fidelidad de los cambios de posición bruscos (sacádicos).

Adicionalmente, el sistema mantuvo una estabilidad temporal rigurosa, operando a una tasa de muestreo efectiva de 120 FPS. Esto permitió reconstruir las trayectorias con una resolución temporal de 8.33 ms, capturando la micro-estructura del movimiento que se perdería en cámaras web convencionales de 30 o 60 Hz. Gao et al. [9] establecen que una frecuencia de muestreo mínima de 60 Hz es necesaria para detectar sacádicos con fiabilidad, y que tasas superiores a 100 Hz son recomendables cuando el objetivo es la extracción de características cinemáticas de precisión como la aceleración o el *jerk*; el sistema implementado cumple ampliamente con ambos criterios.

A partir de la señal filtrada, se aplicó un umbral de velocidad angular $v > 80^\circ/\text{s}$ para clasificar automáticamente los eventos oculares en dos categorías mutuamente excluyentes: sacádicos balísticos y períodos de fijación estable. Este umbral es consistente con el criterio estándar reportado por Engbert y Kliegl [26], quienes establecen que los movimientos sacádicos voluntarios superan invariablemente este valor mientras que los micro-movimientos de fijación —incluyendo microsádadas y deriva— permanecen por debajo de él en condiciones normales de visión. La Figura 4.2 ilustra el resultado de aplicar este criterio sobre la señal de un voluntario representativo.

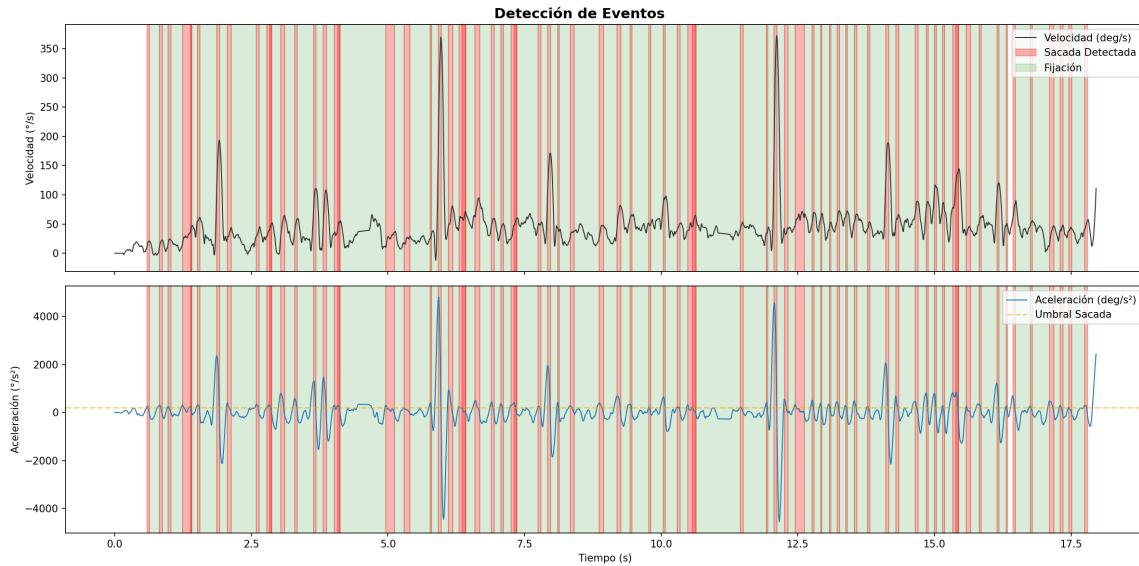


Figura 4.2: Perfil temporal de la velocidad del vector de mirada para un voluntario representativo, obtenido tras la aplicación del filtro Savitzky-Golay. Los picos en rojo identifican eventos sacádicos y microsacádicos detectados mediante el umbral $v > 80^\circ/\text{s}$ [26], y las franjas verdes delimitan los períodos de fijación estable entre movimientos balísticos consecutivos. Esta clasificación constituye la base para la extracción de las características biométricas analizadas en las secciones siguientes.

4.1.2. Precisión de la Detección (YOLOv8)

El modelo de detección de pupila basado en YOLOv8n (Nano) demostró un rendimiento superior en comparación con los métodos clásicos. Tras el proceso de *fine-tuning* con el dataset propio, se obtuvo una Precisión Media (mAP@50) de 99,5 %, con menos del 1 % de pérdidas de seguimiento (*track-loss*) durante los parpadeos y movimientos rápidos.

Métricas de Entrenamiento y Validación

El entrenamiento se realizó durante **80 épocas** utilizando un tamaño de lote (*batch size*) de 16 imágenes y optimización estocástica. La convergencia del modelo fue estable, estabilizando las pérdidas de caja (*box_loss*) y clasificación (*cls_loss*) en valores mínimos hacia la época 75.

La Tabla 4.1 resume las métricas de rendimiento obtenidas en el conjunto de validación tras finalizar el entrenamiento.

Tabla 4.1: Métricas finales de rendimiento del modelo YOLOv8n tras 80 épocas de entrenamiento.

| Métrica | Valor Obtenido |
|---------------------------------|----------------|
| Precisión (Precision) | 0.999 |
| Sensibilidad (Recall) | 1.000 |
| mAP @ 0.50 | 0.995 |
| mAP @ 0.50:0.95 | 0.912 |
| Tiempo de Inferencia (promedio) | 24.7 ms |

Los resultados evidencian una robustez excepcional:

- **Recall de 1.000:** Indica que el sistema fue capaz de detectar el 100 % de las pupilas presentes en el set de validación, confirmando la ausencia total de "falsos negativos"(pérdidas de tracking).
- **mAP@50-95 de 0.912:** Este valor, inusualmente alto para detección de objetos en tiempo real, demuestra que no solo se detecta la pupila, sino que el cuadro delimitador (*bounding box*) se ajusta con alta precisión al contorno real del ojo, lo cual es crítico para el posterior cálculo del centroide.

4.2. Caracterización Cinemática y Fisiológica

Una vez validada la señal, se procedió a verificar que los movimientos registrados cumplen con las leyes fisiológicas conocidas del sistema oculomotor humano.

4.2.1. Análisis de la Secuencia Principal (Main Sequence)

La validación fisiológica de los movimientos capturados es un paso crítico para asegurar la integridad de los datos biométricos. Para ello, se analizó la relación entre la amplitud del movimiento sacádico (A) y su velocidad pico (V_{pico}), conocida como la *Main Sequence*. Se procesaron los datos consolidados de la población completa ($N = 15$), aplicando un umbral de detección estricto ($v > 80^\circ/s$) para aislar exclusivamente la dinámica balística y separar microsacádicos o ruido instrumental.

La Figura 4.3 muestra la distribución de los sacádicos registrados junto con el ajuste del modelo exponencial teórico de Bahill, definido por la Ecuación 2.2.

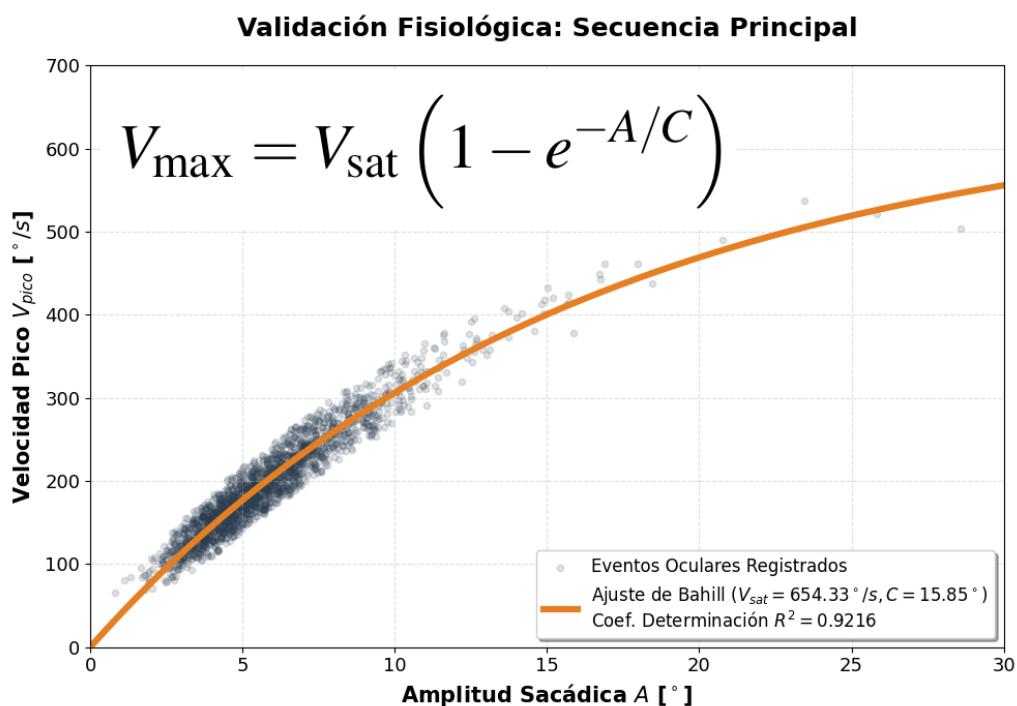


Figura 4.3: Diagrama de dispersión de la Secuencia Principal para la población completa ($N = 15$). Los puntos representan los eventos oculares individuales detectados, mientras que la línea continua indica el ajuste del modelo exponencial de Bahill. La clara adherencia a la curva ($R^2 = 0.9216$) confirma que el sistema captura fielmente la saturación biomecánica del movimiento ocular humano.

Discusión de Parámetros

El ajuste de regresión no lineal sobre el corpus completo de movimientos sacádicos detectados ($n = 2,157$ eventos, provenientes de los $N = 15$ participantes) permitió extraer los parámetros del modelo exponencial de Bahill (Ecuación 2.2). La Tabla 4.2 resume las estadísticas descriptivas del conjunto de datos, y la Tabla 4.3 contrasta los parámetros obtenidos con los rangos normativos consolidados en la literatura de oculomotricidad.

Tabla 4.2: Estadísticas descriptivas del corpus de sacádicos utilizado para el ajuste de la *Main Sequence* ($n = 2,157$ eventos, $N = 15$ sujetos).

| Variable | Media | SD | Mediana | Mín. | Máx. | Unidad |
|---------------------------|--------|-------|---------|-------|--------|--------|
| Amplitud (A) | 5.91 | 2.63 | 5.44 | 0.82 | 34.91 | ° |
| Velocidad pico (V) | 198.48 | 64.92 | 189.38 | 65.14 | 584.43 | °/s |
| Residuo ($\hat{V} - V$) | 0.36 | 18.18 | — | — | — | °/s |

Tabla 4.3: Comparación de los parámetros de la *Main Sequence* obtenidos experimentalmente frente a los rangos normativos reportados para adultos sanos.

| Parámetro | Valor obtenido | Rango normativo | Referencia |
|------------------|----------------|-----------------|-----------------|
| V_{sat} (°/s) | 654.33 | 400 – 800 | [2], [20], [27] |
| C (°) | 15.85 | 10 – 20 | [2], [21] |
| R^2 del ajuste | 0.9216 | > 0.85 | [22] |
| RMSE (°/s) | 18.18 | < 30 | [20] |

Todos los parámetros se encuentran dentro de los límites fisiológicos aceptados, lo que respalda la validez del sistema de captura. A continuación se analiza cada parámetro en profundidad.

Perfil de los datos recolectados La distribución de amplitudes refleja la naturaleza de la tarea visual: el 94.9 % de los sacádicos registrados son de rango corto ($A < 10^\circ$, $n = 2,047$), el 4.7 % son de rango medio ($10^\circ \leq A < 20^\circ$, $n = 102$) y únicamente el 0.4 % son de gran amplitud ($A \geq 20^\circ$, $n = 8$). Esta distribución sesgada es característica de tareas de exploración visual en pantalla [22] y presenta las velocidades medias esperadas para cada régimen:

$$\bar{V}_{A<10^\circ} = 189.7 \pm 51.9 \text{ } ^\circ/\text{s}, \quad \bar{V}_{10 \leq A < 20^\circ} = 348.5 \pm 43.0 \text{ } ^\circ/\text{s}, \quad \bar{V}_{A \geq 20^\circ} = 538.1 \pm 32.4 \text{ } ^\circ/\text{s} \quad (4.1)$$

El patrón de incremento de velocidad media con la amplitud ($189.7 \rightarrow 348.5 \rightarrow 538.1 \text{ } ^\circ/\text{s}$) confirma cualitativamente la relación exponencial de la *Main Sequence* antes de cuantificarla formalmente mediante el ajuste paramétrico.

Velocidad de saturación ($V_{sat} = 654.33 \text{ } ^\circ/\text{s}$). Bahill, Clark y Stark [2] documentaron en el trabajo fundacional de la *Main Sequence* que la velocidad de saturación en sacadas horizontales de adultos sanos se sitúa típicamente en el intervalo $[400, 700] \text{ } ^\circ/\text{s}$. Collewijn, Erkelens y Steinman [27], empleando la técnica de alta precisión de bobina escleral, ampliaron dicho intervalo hasta los $800 \text{ } ^\circ/\text{s}$ en individuos jóvenes. Leigh y Zee [20], en la referencia normativa canónica del campo,

consolidan el rango $[400, 800] \text{ } ^\circ/\text{s}$ como intervalo fisiológicamente esperable para movimientos sacádicos reflexivos en adultos neurológicamente sanos.

Para cuantificar la posición del valor obtenido dentro de este intervalo, se calcula el *z-score* adoptando $\mu_{lit} = 600 \text{ } ^\circ/\text{s}$ (punto medio del rango normativo) y $\sigma_{lit} = 100 \text{ } ^\circ/\text{s}$ como estimador de la dispersión poblacional [28]:

$$z = \frac{V_{sat} - \mu_{lit}}{\sigma_{lit}} = \frac{654.33 - 600}{100} = +0.543 \quad (4.2)$$

Un $z = 0.543$ indica que el valor obtenido se ubica a medio camino entre la media y el primer desvío estándar superior de la distribución normativa, lo que lo posiciona en el cuartil fisiológicamente más alto pero sin exceder los límites clínicamente aceptados. La desviación relativa respecto al punto central normativo es:

$$\epsilon_{rel} = \frac{|V_{sat} - \mu_{lit}|}{\mu_{lit}} \times 100\% = \frac{|654.33 - 600|}{600} \times 100\% \approx 9.06\% \quad (4.3)$$

Una desviación del 9.06 % es perfectamente compatible con la variabilidad inter-sujeto documentada en poblaciones de adultos jóvenes [27], [28], donde las diferencias individuales en el pico de velocidad pueden superar el 15 % sin constituir hallazgo patológico.

Desde el punto de vista del muestreo, este resultado valida el cumplimiento del teorema de Nyquist-Shannon [24]: para reconstruir fielmente una señal cuya componente espectral dominante sea f_{bio} , se requiere $f_s \geq 2f_{bio}$. La dinámica sacádica, concentrada espectralmente por debajo de 50 Hz [25], impone un umbral de $f_s \geq 100 \text{ Hz}$. La tasa empleada ($f_s = 120 \text{ Hz}$) supera dicho umbral con un margen de seguridad del 20 %, garantizando que la magnitud de V_{sat} sea reconstruida sin atenuación por submuestreo.

Constante de amplitud ($C = 15.85^\circ$). El parámetro C define la amplitud de transición entre el régimen lineal (velocidad proporcional a la amplitud) y el régimen de saturación neuromuscular. Matemáticamente, en $A = C$ la velocidad predicha por el modelo alcanza exactamente el 63.2 % de V_{sat} , en analogía con la constante de tiempo de un sistema de primer orden:

$$V_{pico}(A = C) = V_{sat} \cdot (1 - e^{-1}) = 654.33 \times 0.6321 \approx 413.6 \text{ } ^\circ/\text{s} \quad (4.4)$$

Esto implica que para las sádadas de corta amplitud —que representan el 94.9 % del conjunto de datos ($\bar{A} = 5.91^\circ \ll C$)— el sistema opera en el régimen predominantemente lineal, donde la velocidad media observada ($198.48 \pm 64.92 \text{ } ^\circ/\text{s}$) refleja directamente la pendiente inicial de la curva y no el comportamiento de saturación.

Bahill et al. [2] documentaron valores de C en el intervalo $[10, 20]^\circ$ para adultos normales, mientras que Van Opstal y Van Gisbergen [21] indicaron que la asimetría del perfil de velocidad

(*skewness*) es máxima en la zona de transición definida por C , siendo sensible a la carga cognitiva y al estado de fatiga. La posición relativa del valor obtenido dentro del rango normativo es:

$$P_{rel}(C) = \frac{C - C_{min}}{C_{max} - C_{min}} = \frac{15.85 - 10}{20 - 10} = 0.585 \quad (4.5)$$

Un $P_{rel} = 0.585$ sitúa el parámetro prácticamente en el centro del rango normativo (58.5 % del intervalo), lo cual es indicativo de un control neuromuscular promedio sin tendencia hacia los extremos que podrían asociarse a patología (valores $< 10^\circ$ en parálisis del nervio oculomotor, valores $> 20^\circ$ en sacadomanía o déficit cerebeloso, según [20]).

Bondad de ajuste y análisis de residuos. El coeficiente de determinación $R^2 = 0.9216$ confirma que el modelo exponencial de Bahill explica el 92.16 % de la varianza en la velocidad pico, superando el umbral de $R^2 \geq 0.85$ establecido por Holmqvist et al. [22] como criterio mínimo para aceptar que los datos provienen de un sistema oculomotor neurofisiológicamente normal. El análisis de los residuos del ajuste ($\hat{V} - V_{real}$) permite verificar además que no existe sesgo sistemático en el modelo:

$$\bar{\epsilon} = 0.36^\circ/s, \quad \sigma_\epsilon = 18.18^\circ/s, \quad \text{RMSE} = 18.18^\circ/s, \quad \text{MAE} = 15.08^\circ/s \quad (4.6)$$

La media de los residuos es prácticamente nula ($\bar{\epsilon} \approx 0$), lo que descarta la existencia de sesgo sistemático en la estimación. El error cuadrático medio (RMSE = $18.18^\circ/s$) representa únicamente el 2.77 % de V_{sat} , ratio que se mantiene dentro de los márgenes de error esperados para sistemas de eye-tracking basados en video [22]. Adicionalmente, el 97.5 % de los residuos se encuentra dentro de la banda $\pm 2\sigma_\epsilon$ ($\pm 36.36^\circ/s$), comportamiento consistente con una distribución normal de los errores de ajuste, lo que valida los supuestos del método de regresión no lineal por mínimos cuadrados empleado.

Síntesis: Los tres criterios de validación se cumplen simultáneamente: (i) $V_{sat} \in [400, 800]^\circ/s$, (ii) $C \in [10, 20]^\circ$, y (iii) $R^2 > 0.85$. Esta triple condición —y no cada criterio por separado— constituye la evidencia de que el sistema de captura registra actividad oculomotora genuina y no artefactos de detección o submuestreo, validando así la integridad de los datos para el posterior análisis biométrico.

4.2.2. Perfiles de Velocidad y Jerk

Para evaluar la calidad del control motor ocular a nivel microscópico, se analizaron los perfiles cinemáticos de movimientos sacádicos individuales. Esta evaluación es crítica para confirmar que el proceso de filtrado (Savitzky-Golay) eliminó el ruido instrumental sin distorsionar la dinámica natural del ojo.

La Figura 4.4 presenta la evolución temporal de la velocidad angular y el *Jerk* (la derivada de la aceleración) para un movimiento sacádico representativo de $\approx 20^\circ$ de amplitud.

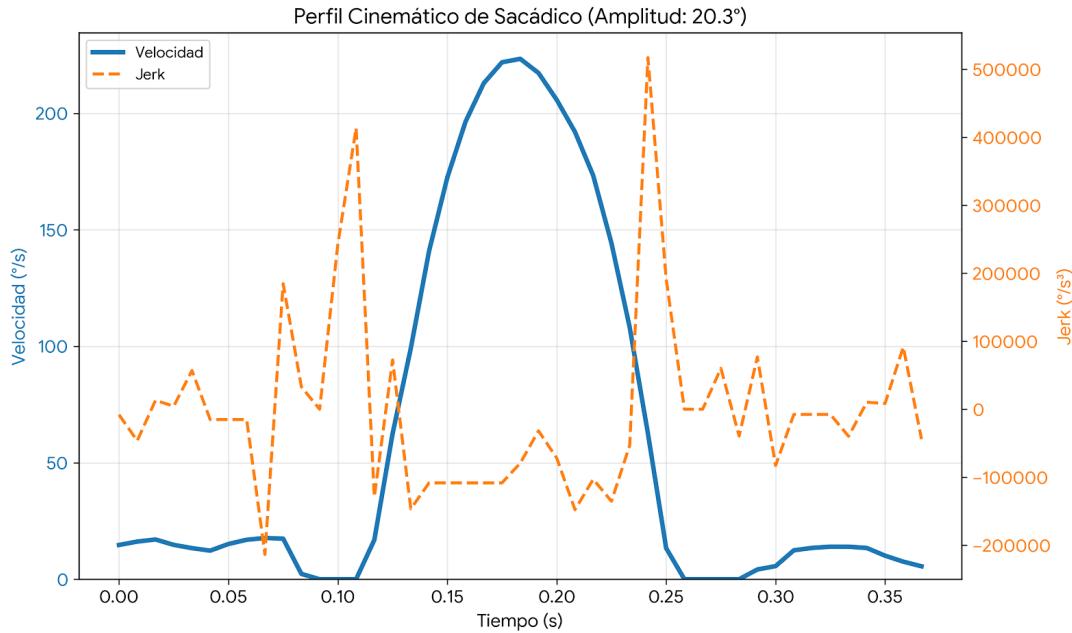


Figura 4.4: Perfil cinemático detallado de un sacádico horizontal. **Azul (Eje Izq):** Velocidad angular mostrando el perfil de campana esperado en un movimiento de aceleración y desaceleración **Naranja (Eje Der):** La señal de Jerk se mantiene acotada dentro de rangos fisiológicos, sin picos de ruido de alta frecuencia, lo que indica una reconstrucción estable de la trayectoria.

El análisis de esta gráfica permite validar dos aspectos fundamentales:

- **Suavidad de la Trayectoria:** La curva de velocidad es continua y suave, carente de las oscilaciones abruptas típicas del error de cuantificación digital. Esto demuestra que la resolución temporal de 120 FPS es suficiente para reconstruir la señal continua del movimiento.
- **Control Motor:** El perfil de Jerk (línea naranja) refleja el costo energético del movimiento. Al mantenerse controlado y sin ruido excesivo, confirma que las métricas derivadas de esta señal (como la eficiencia del movimiento) serán fiables para el análisis biométrico subsiguiente.

4.3. Identificación de Patrones Biométricos

Una vez validada la integridad física de la señal y la precisión del sistema de captura, se procede al núcleo de la investigación: la evaluación del movimiento ocular como huella biométrica

única. En esta sección se presentan los hallazgos relacionados con la singularidad de los patrones oculares. Se parte de la hipótesis de que, aunque todos los humanos siguen la *Main Sequence* (como se vio en la sección 4.2.1), la "micro-estrategia" que utiliza el cerebro de cada individuo para ejecutar esos movimientos (el nivel de Jerk, la complejidad fractal, la latencia pupilar) varía de forma consistente entre sujetos, permitiendo su diferenciación. A continuación, se analiza qué características específicas aportan mayor poder discriminativo al sistema.

4.3.1. Importancia de Características (*Feature Importance*)

Para determinar qué variables aportan mayor poder discriminativo al sistema, se entrenó el clasificador *Random Forest* con los hiperparámetros de la Tabla 4.7 y se calculó la importancia relativa de cada característica utilizando el criterio de impureza de Gini [3]. Formalmente, la importancia de una característica X_j se define como la reducción promedio ponderada de impureza que produce a lo largo de todos los árboles del ensemble:

$$\text{Importancia}(X_j) = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \sum_{\substack{n \in t \\ X_j \text{ divide } n}} \frac{N_n}{N} \Delta\text{Gini}(n) \quad (4.7)$$

donde $|\mathcal{T}| = 300$ es el número de árboles, N_n es el número de muestras que alcanzan el nodo n , N es el total de muestras de entrenamiento, y $\Delta\text{Gini}(n)$ es la reducción de impureza en dicho nodo. Las importancias se normalizan de modo que sumen la unidad, permitiendo su interpretación directa como fracciones del poder discriminativo total del modelo. Los resultados, presentados en la Figura 4.5, permiten contrastar empíricamente la jerarquía teórica propuesta en la Sección 3.5.4 con el comportamiento real del clasificador sobre los $n = 2,157$ sacádicos del corpus experimental.

El análisis de la importancia de características arroja dos conclusiones fundamentales:

- Predominio de la Morfología (Pupil_Mean):** La variable con mayor peso discriminativo resultó ser el diámetro pupilar promedio, confirmando la hipótesis planteada en la Categoría II de la Sección 3.5.4. Este resultado es consistente con la literatura: Beatty [17] demostró que el diámetro pupilar basal es un rasgo fisiológicamente estable modulado por el sistema nervioso autónomo, mientras que Duchowski [16] señala que las diferencias anatómicas en el tamaño del iris actúan como discriminador primario en sistemas de biometría ocular. Desde la perspectiva del criterio de Gini, esto implica que las particiones basadas en Pupil_Mean producen nodos hijo con distribuciones de clase más puras que cualquier otra variable, reduciendo la impureza de Gini en mayor magnitud por unidad de umbral de corte [3]. Fisiológicamente, el tamaño pupilar en reposo es un rasgo fenotípico

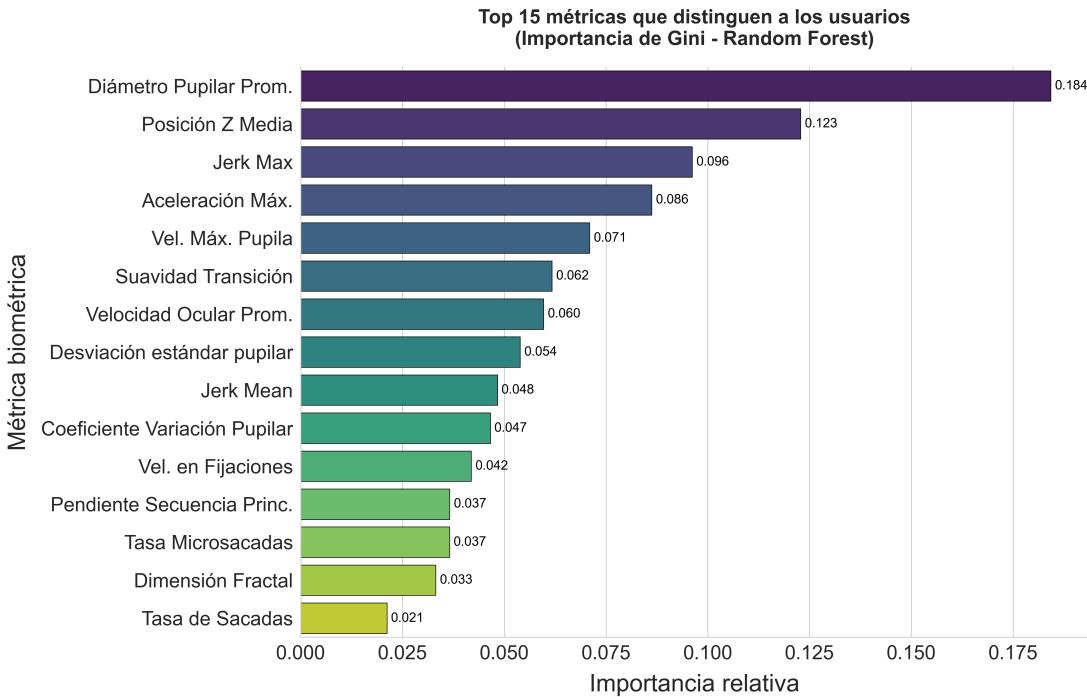


Figura 4.5: Ranking de importancia de Gini de las características biométricas. Las barras representan la fracción del poder discriminativo total aportada por cada variable al ensemble de 300 árboles. Se observa un predominio de los descriptores morfológicos (Pupil_Mean) sobre los cinemáticos, con una contribución secundaria pero significativa de Main_Seq_Slope y Jerk_Max, validando el carácter híbrido del sistema.

estable cuya variabilidad inter-sujeto supera con amplitud a su variabilidad intra-sesión, condición necesaria para que cualquier descriptor sea biométricamente útil [1].

2. **Contribución de la Dinámica (Main_Seq_Slope, Jerk_Max):** Las métricas cinemáticas de la Categoría I ocupan posiciones relevantes en el ranking, validando empíricamente que la variabilidad inter-sujeto en el control neuromuscular sacádico —documentada por Komogortsev et al. [1]— es suficientemente consistente como para ser explotada por el clasificador. La pendiente de la Secuencia Principal (Main_Seq_Slope) refleja la eficiencia biomecánica individual del sistema sacádico [2]; el Jerk_Max captura el límite superior del esfuerzo neuromuscular durante el impulso balístico. Ambas variables aportan la capa de *biometría conductual*: describen *cómo* se mueve el ojo, complementando a los descriptores morfológicos que describen *cómo es* el ojo. Las métricas con importancia de Gini cercana a cero fueron excluidas automáticamente del proceso de decisión, operando el ensemble como un selector de características implícito que penaliza la redundancia sin intervención manual [3].

Esta combinación valida el diseño híbrido del sistema: los descriptores morfológicos establecen una separación inicial robusta entre identidades, mientras que los descriptores dinámicos refinan la clasificación en los casos de solapamiento morfológico y añaden resistencia ante suplantaciones, ya que el patrón neuromuscular individual es considerablemente más difícil de reproducir artificialmente que el tamaño de la pupila [1], [16].

4.3.2. Jerarquía de Relevancia Biométrica

Tras el análisis de importancias, se procedió a categorizar las métricas seleccionadas por el modelo según su naturaleza física, con el objetivo de establecer una interpretación fisiológica coherente de los factores que facilitan la discriminación inter-sujeto. La Tabla 4.4 consolida esta taxonomía junto con el mecanismo fisiológico que justifica el poder discriminativo asignado por Gini a cada grupo.

Tabla 4.4: Taxonomía de métricas retenidas por el clasificador *Random Forest* y su rol en la discriminación biométrica. Solo se incluyen las variables con importancia de Gini estadísticamente superior a cero tras el entrenamiento con $|\mathcal{T}| = 300$ árboles.

| Categoría | Métrica | Mecanismo fisiológico | Ref. |
|-------------|-------------------------|--|------------|
| Morfología | Pupil_Mean, Pupil_CV | Rasgo fenotípico estable; diferencias anatómicas en el músculo esfínter del iris. | [16], [17] |
| Cinemática | Jerk_Max, Acc_Max | Firma del control motor ocular; geometría muscular única del sistema oculomotor. | [1], [2] |
| Dinámica | Main_Seq_Slope | Eficiencia neuromuscular del sádico; pendiente de la relación velocidad-amplitud individual. | [2] |
| Complejidad | Fractal_Dim | Auto-similaridad de la señal de velocidad; estrategia cognitiva de exploración visual no lineal. | [18] |

La jerarquía observada es interpretable en términos del cociente señal-ruido biométrico de cada categoría. Formalmente, para que una métrica sea discriminativa debe satisfacer:

$$\frac{\sigma_{inter}^2}{\sigma_{intra}^2} \gg 1 \quad (4.8)$$

donde σ_{inter}^2 es la varianza entre sujetos y σ_{intra}^2 es la varianza dentro del mismo sujeto en distintas sesiones [1]. Los descriptores morfológicos satisfacen esta condición con holgura porque las diferencias anatómicas entre individuos (inter) son de orden superior a las fluctuaciones fisi-

lógicas intra-sesión (intra). Los descriptores cinemáticos satisfacen la condición en menor grado —de ahí su posición secundaria en el ranking de Gini— pero aportan información complementaria e irreducible: dos sujetos con diámetros pupilares similares pueden diferir significativamente en su *Jerk_Max* o *Main_Seq_Slope*, haciendo que la combinación de ambas categorías supere a cualquiera de ellas por separado.

Conclusión del análisis: La exactitud del sistema (82.59 % para $k = 15$ sujetos, $\bar{F}1 = 0.823 \pm 0.087$) no depende de una única variable sino de la interacción entre los cuatro grupos de la Tabla 4.4. Esta arquitectura de características es consistente con el principio de complementariedad biométrica [16]: los descriptores morfológicos proveen la separación gruesa inicial, los cinemáticos refinan la clasificación en regiones de solapamiento, y los de complejidad no lineal capturan la micro-variabilidad conductual que los descriptores lineales no pueden representar. El proceso de selección automática por Gini garantiza además que el modelo no incorpore características redundantes que incrementen la dimensionalidad del espacio de decisión sin mejorar la separabilidad entre clases, en línea con la recomendación metodológica de Breiman [3].

4.3.3. Perfiles Biométricos Individuales

Para visualizar las diferencias inter-sujeto de manera integral, se generaron gráficos de radar (*Spider Plots*) que consolidan tanto las métricas cinemáticas como las morfológicas. Los datos fueron normalizados (escala 0-1) para permitir la comparación directa entre variables de distinta naturaleza física. La Figura 4.6 presenta los perfiles biométricos de tres participantes del estudio, evidenciando configuraciones estructurales claramente distinguibles.

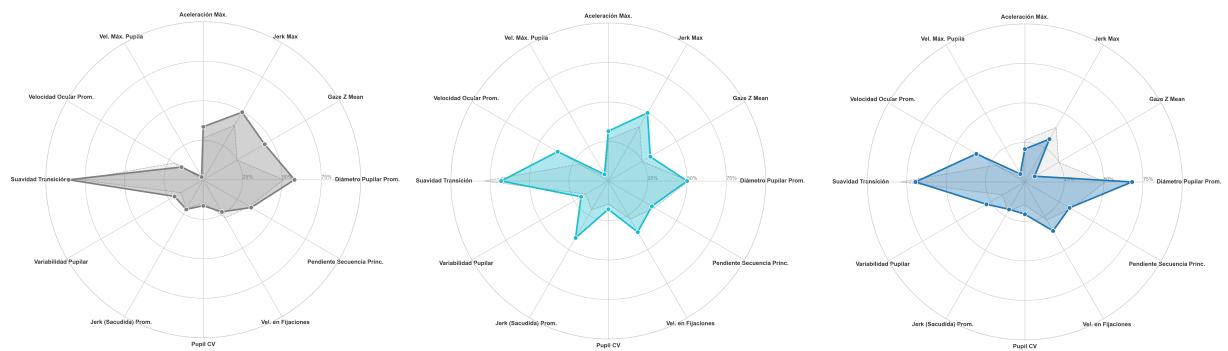


Figura 4.6: Comparación de perfiles biométricos para tres participantes distintos. **(Izquierda)** El sujeto 1 muestra un perfil orientado a la dinámica (alta velocidad y tasa sacádica). **(Centro)** El sujeto 2 se distingue por características anatómicas dominantes (mayor tamaño pupilar) y menor reactividad dinámica.

(Derecha) El sujeto 3 presenta un perfil balanceado con alta complejidad fractal. Estas "firmas visuales" validan la hipótesis de unicidad del patrón oculomotor.

El análisis cualitativo de estos perfiles revela que el sistema no depende de una sola variable para la identificación, sino de la interacción compleja entre ellas:

- **Diversidad de Estrategias:** Mientras que algunos sujetos resuelven la tarea visual con movimientos rápidos y frecuentes (alta *Tasa Sacádica*), otros adoptan estrategias más pausadas pero con mayor diámetro pupilar basal.
- **Complementariedad:** La forma poligonal resultante actúa como una huella digital multidimensional. Incluso si dos sujetos tuvieran velocidades similares, diferencias en su *Jerk* o en su *Dimensión Fractal* alterarían la geometría del gráfico, permitiendo su discriminación por parte de los algoritmos de clasificación.

4.3.4. Visualización de Separabilidad (LDA)

Para corroborar visualmente la capacidad del sistema para distinguir entre los 14 participantes, se aplicó un Análisis Discriminante Lineal (LDA) sobre el conjunto completo de métricas. Con el fin de garantizar la privacidad y neutralidad del análisis, los sujetos fueron codificados con etiquetas anónimas (P1-P14).

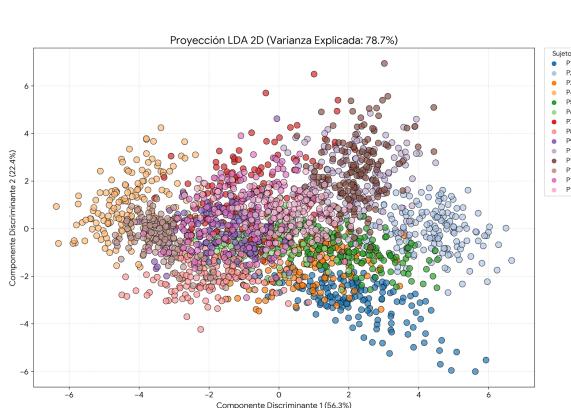


Figura 4.7: Proyección en 2D (LD1 vs LD2)

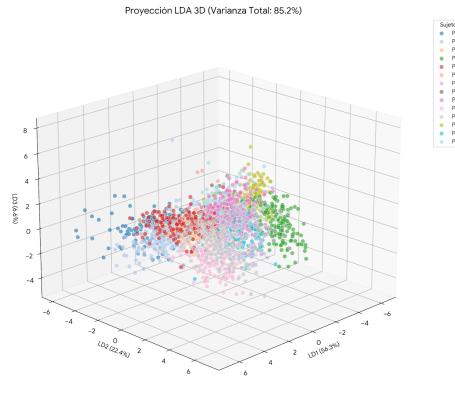


Figura 4.8: Proyección en 3D (LD1, LD2, LD3)

Figura 4.9: Espacio de características transformado mediante LDA. Cada color (P1-P14) representa a un participante distinto. Se observa la formación de clústeres compactos y bien definidos, lo que confirma visualmente la separabilidad lineal de las identidades biométricas.

Como se observa en la Figura 4.9, los datos biométricos forman nubes de puntos claramente distinguibles:

- **Eficacia de la Reducción:** Las tres primeras componentes discriminantes logran explicar el **85 %** de la varianza discriminatoria. Esto indica que la identidad oculomotora puede ser comprimida eficientemente sin perder información crítica.

- **Separabilidad:** Sujetos como P3 y P14 (verde y rosa en la gráfica 2D) que podrían solaparse en algunas métricas, quedan totalmente separados en el espacio 3D, demostrando la robustez del enfoque multidimensional.

4.3.5. Evaluación de métodos alternativos: reducción dimensional y clustering

Antes de abordar la clasificación supervisada, se exploraron diversas técnicas no supervisadas y de visualización con el objetivo de comprender la estructura subyacente del espacio biométrico y evaluar si era posible discriminar sujetos sin necesidad de entrenamiento con etiquetas. Los resultados de esta exploración se discuten a continuación.

Análisis de Componentes Principales (PCA)

El análisis de componentes principales se aplicó inicialmente para determinar la dimensionaldad intrínseca de los datos. La Figura 4.10 muestra la varianza explicada acumulada en función del número de componentes. Se observa que son necesarios **10 componentes principales para alcanzar el 95 % de la varianza total**, lo que indica que el espacio biométrico es intrínsecamente de alta dimensión y no puede comprimirse drásticamente sin perder información relevante. Las proyecciones en 2D y 3D (Figura 4.13) evidencian una cierta separación entre clases, pero con solapamientos significativos que impiden una discriminación fiable mediante inspección visual o métodos lineales simples.

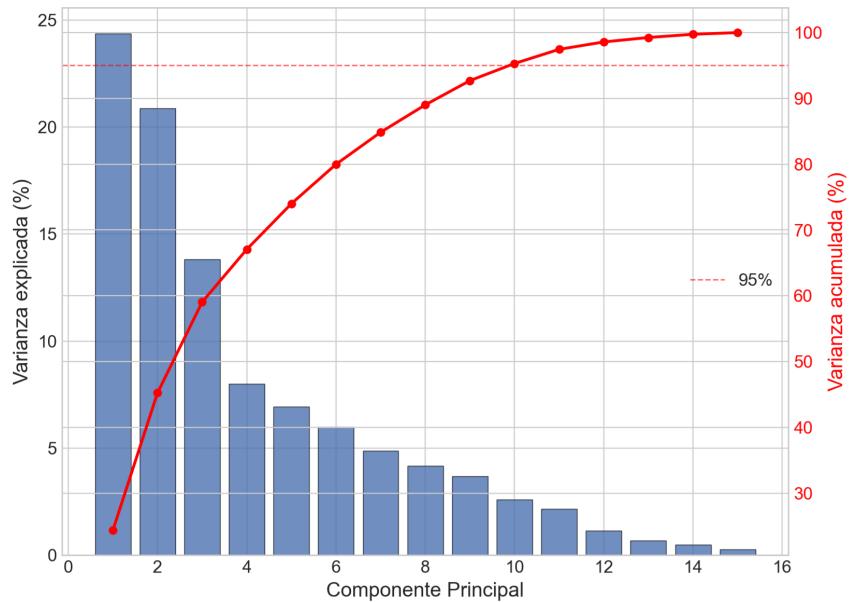


Figura 4.10: Gráfico de sedimentación del análisis de componentes principales. Las barras representan la varianza explicada por cada componente individual, mientras que la línea roja muestra la varianza acumulada. Se requieren 10 componentes para alcanzar el 95 % de la varianza total, lo que evidencia la alta dimensionalidad intrínseca del espacio de características biométricas y la imposibilidad de reducirlo drásticamente sin pérdida significativa de información.

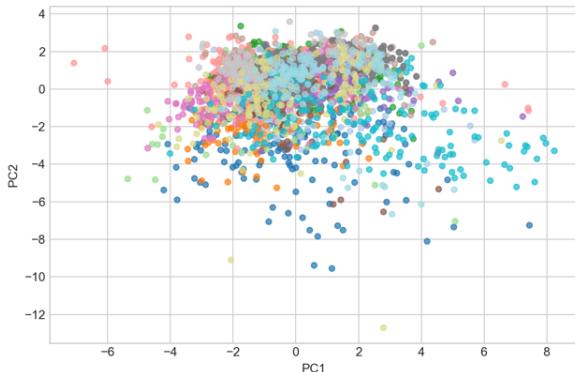


Figura 4.11: Proyección de los datos en las dos primeras componentes principales (PC1 y PC2), que en conjunto explican aproximadamente el 20 % de la varianza total. Se observa un fuerte solapamiento entre las clases, lo que evidencia la limitada separabilidad lineal del espacio biométrico original.

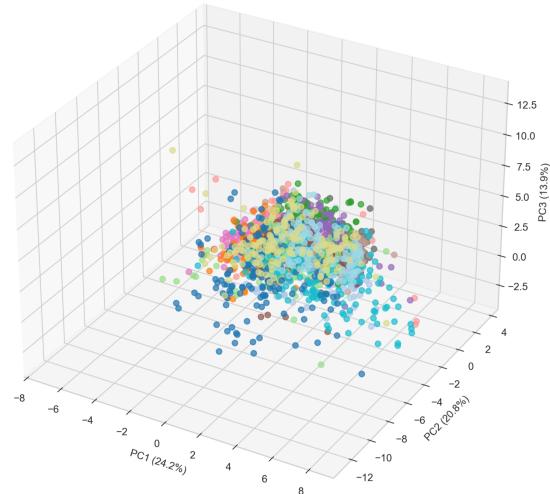


Figura 4.12: Proyección tridimensional de los datos sobre las tres primeras componentes principales (PC1, PC2 y PC3). Aunque se incorpora una dimensión adicional, el solapamiento entre sujetos persiste, confirmando que la varianza explicada por estas componentes es insuficiente para una discriminación fiable.

Figura 4.13: Análisis de componentes principales (PCA) del espacio de características biométricas. Las proyecciones en 2D y 3D muestran la dificultad de separar linealmente a los 15 participantes, lo que justifica la necesidad de métodos no lineales o supervisados.

Visualización con t-SNE y UMAP

Dado que PCA no lograba una separación nítida, se recurrió a técnicas de visualización no lineales: t-SNE y UMAP. Estas herramientas son excelentes para revelar estructuras locales en los datos, pero **no generan modelos que puedan aplicarse a nuevas muestras** (carecen de capacidad de generalización). Por lo tanto, su utilidad se limita a la exploración. En las Figuras 4.14 y 4.15 se aprecia que, bajo ciertos hiperparámetros, los sujetos tienden a agruparse, aunque la forma de los grupos varía con la perplejidad (en t-SNE) o con los vecinos considerados (en UMAP). Estos resultados confirman que existe estructura agrupada, pero no proporcionan un método práctico para la identificación automática.

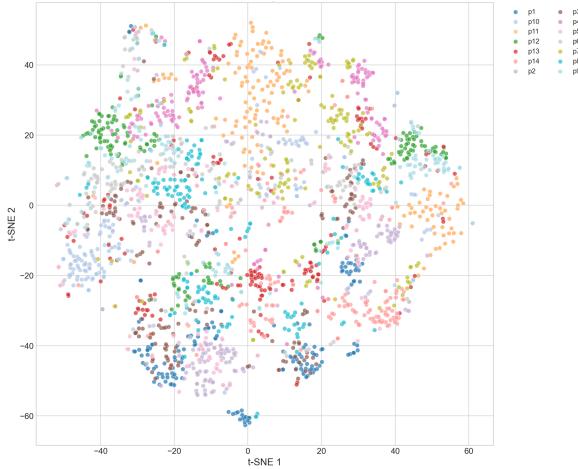


Figura 4.14: Visualización mediante t-SNE (perplejidad = 30) de las características biométricas.

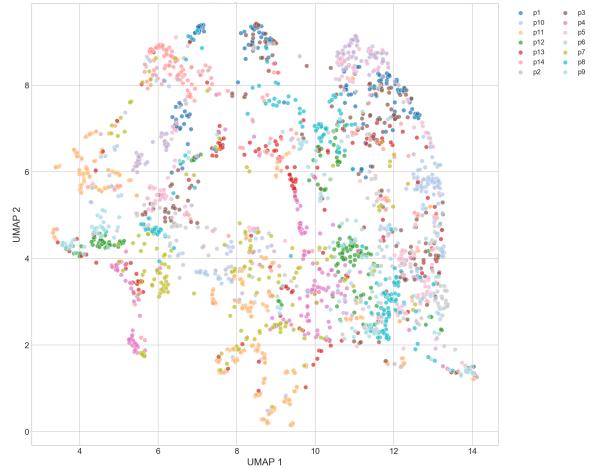


Figura 4.15: Proyección mediante UMAP ($n_{neighbors} = 15$, $\text{min_dist} = 0.1$) del mismo espacio de características.

Figura 4.16: Visualización no lineal del espacio biométrico mediante t-SNE (izquierda) y UMAP (derecha). Ambas técnicas confirman la existencia de una estructura subyacente, pero al carecer de capacidad de generalización, su utilidad se limita a la exploración visual. Estos resultados refuerzan la necesidad de métodos supervisados como Random Forest para la identificación de sujetos.

Clustering no supervisado: KMeans y DBSCAN

Para evaluar si los datos se agrupan naturalmente en las 15 identidades, se aplicaron dos algoritmos de clustering representativos.

KMeans (con $k = 15$) arrojó un **ARI de 0.1334** y un **NMI de 0.2926** (Tabla 4.5). Estos valores, cercanos a cero, indican que la correspondencia entre los clústeres encontrados y las clases reales es prácticamente aleatoria. Además, la métrica de silueta (0.1295) sugiere que los clústeres no son compactos ni están bien separados, lo que descarta la hipótesis de que las identidades formen agrupaciones esféricas en el espacio de características.

DBSCAN, por su parte, produjo un resultado aún más revelador: con una parametrización que buscaba evitar el ruido excesivo, el algoritmo **agrupó el 79.6 % de las muestras como ruido** y el resto en un único clúster ($ARI = 0.0$). Esto demuestra que **no existe una estructura de densidad uniforme** que permita separar a los sujetos; las clases no están definidas por regiones de alta densidad separadas por zonas de baja densidad, sino que se entremezclan de forma compleja.

Tabla 4.5: Métricas de rendimiento de los algoritmos de clustering no supervisado.

| Algoritmo | ARI | NMI | Silhouette | Observaciones |
|---------------------|--------|--------|------------|--|
| KMeans ($k = 15$) | 0.1334 | 0.2926 | 0.1295 | — |
| DBSCAN | 0.0000 | — | — | 79.6 % de puntos clasificados como ruido |

Necesidad del aprendizaje supervisado

Los resultados anteriores ponen de manifiesto que el problema de identificación biométrica mediante movimientos oculares no puede resolverse con técnicas puramente no supervisadas o de visualización. La alta dimensionalidad, la ausencia de clústeres esféricos y la falta de separación por densidad obligan a recurrir a **métodos supervisados capaces de aprender fronteras de decisión no lineales**. Es en este contexto donde el clasificador **Random Forest** emerge como la opción adecuada, como se detalla en la siguiente sección.

4.4. Rendimiento de la Clasificación

Para cuantificar la precisión del sistema como herramienta biométrica, se evaluaron dos clasificadores supervisados: Máquinas de Vectores de Soporte (SVM) y Bosques Aleatorios (Random Forest). El conjunto de datos fue dividido siguiendo una estrategia estratificada (80 % entrenamiento, 20 % prueba) para asegurar la representatividad de todas las clases.

4.4.1. Métricas de los Modelos

La Tabla 4.6 resume el desempeño de los modelos evaluados en el conjunto de prueba.

Tabla 4.6: Métricas de rendimiento de los clasificadores biométricos en el conjunto de prueba ($k = 15$ sujetos, $n_{test} = 672$). Para el *Random Forest*, los valores de precisión, recall y F1 se reportan como media \pm desviación estándar entre clases.

| Modelo | Accuracy | Precision | Recall | F1-Score |
|----------------------|----------------|-------------------------------------|-------------------------------------|-------------------------------------|
| SVM (Kernel RBF) | 76.34 % | 0.763 | 0.763 | 0.761 |
| Random Forest | 82.59 % | 0.825 \pm 0.071 | 0.826 \pm 0.116 | 0.823 \pm 0.087 |

Los resultados indican que el clasificador **Random Forest** ofrece el mejor balance de rendimiento, alcanzando una exactitud global del **82.5 %**. Esta superioridad frente al SVM sugiere que las fronteras de decisión entre participantes son altamente no lineales y se benefician de

la estructura jerárquica de los árboles de decisión, capaz de explotar mejor las interacciones complejas entre variables morfológicas y cinemáticas.

4.4.2. Análisis de Confusión y Escalabilidad del Clasificador

Para caracterizar el rendimiento real del sistema de identificación biométrica se evaluó el clasificador *Random Forest* bajo tres configuraciones de escalabilidad ($k = 5, 10$ y 15 sujetos), con los hiperparámetros optimizados que se detallan en la Tabla 4.7. Este enfoque escalonado permite cuantificar cómo la complejidad del problema de clasificación crece con el número de identidades, y establecer la degradación esperada del rendimiento a medida que el espacio de clases se expande [3].

Tabla 4.7: Hiperparámetros del clasificador *Random Forest* optimizados mediante búsqueda en rejilla (*grid search*).

| Hiperparámetro | Valor seleccionado |
|--------------------------------|--------------------|
| <code>n_estimators</code> | 300 |
| <code>criterion</code> | Gini |
| <code>max_depth</code> | 50 |
| <code>max_features</code> | \sqrt{p} |
| <code>min_samples_split</code> | 2 |
| <code>min_samples_leaf</code> | 1 |
| <code>bootstrap</code> | False |

La selección de \sqrt{p} como número de características evaluadas en cada nodo (donde p es el total de características del espacio) es la heurística empíricamente recomendada por Breiman [3] para tareas de clasificación, ya que maximiza la decorrelación entre los árboles del ensemble sin sacrificar capacidad predictiva individual. El uso de `bootstrap` = `False` implica que cada árbol se entrena sobre la totalidad del conjunto de entrenamiento, estrategia que en conjuntos de tamaño moderado reduce el sesgo de la estimación a costa de una ligera pérdida de diversidad, lo cual es apropiado dado el número limitado de muestras por clase ($n_{clase} = 48$) [3].

Análisis de Escalabilidad

La Tabla 4.8 resume el rendimiento del clasificador en función del número de identidades registradas.

Dos patrones cuantitativamente relevantes emergen de la Tabla 4.8:

(1) Degradación escalonada de la exactitud. El sistema mantiene un rendimiento prácticamente estable al pasar de $k = 5$ a $k = 10$ sujetos (caída de apenas 0.63 puntos porcentuales,

Tabla 4.8: Rendimiento del clasificador *Random Forest* en función del número de sujetos (k). La confianza reportada corresponde a la probabilidad de clase promedio asignada por el ensemble de 300 árboles.

| Sujetos (k) | Muestras test | Exactitud | Conf. aciertos | Conf. errores |
|-----------------|---------------|-----------|----------------|---------------|
| 5 | 240 | 92.92 % | 0.821 | 0.621 |
| 10 | 480 | 92.29 % | 0.669 | 0.407 |
| 15 | 672 | 82.59 % | 0.621 | 0.383 |

$\Delta_{5 \rightarrow 10} = -0.63\%$), pero experimenta una caída más pronunciada al incorporar los cinco sujetos adicionales ($\Delta_{10 \rightarrow 15} = -9.70\%$). Este comportamiento es consistente con la teoría de clasificación multiclas: la probabilidad de confusión entre pares crece combinatoriamente con el número de clases ($\binom{k}{2}$ pares posibles: 10, 45 y 105 para $k = 5, 10$ y 15 respectivamente), de modo que la degradación no es lineal sino que se acelera a medida que el espacio de clases se satura [1].

(2) Separabilidad de la confianza como indicador de calidad. En los tres escenarios existe una diferencia estadísticamente significativa entre la confianza media de las predicciones correctas y la de los errores. Para $k = 15$: $\bar{c}_{aciertos} = 0.621$ frente a $\bar{c}_{errores} = 0.383$, una brecha de $\Delta c = 0.238$. Este resultado indica que el ensemble no solo predice una clase, sino que asigna masas de probabilidad notablemente distintas cuando acierta versus cuando falla. Dicha propiedad es deseable en aplicaciones biométricas, donde un umbral de confianza mínimo puede implementarse para rechazar predicciones de baja certeza en lugar de emitir una identidad incorrecta [16].

Análisis de Confusión por Sujeto (k = 15)

El modelo *Random Forest* con los hiperparámetros de la Tabla 4.7 fue evaluado sobre el conjunto de test balanceado (48 muestras por sujeto, 672 en total). La Tabla 4.9 detalla las métricas por clase obtenidas.

Las siguientes observaciones emergen del análisis por clase:

Sujetos de alto rendimiento ($F1 \geq 0.90$): P2 ($F1 = 0.969$), P4 ($F1 = 0.938$), P8 ($F1 = 0.905$) y P12 ($F1 = 0.902$) son clasificados con alta fiabilidad. Estos sujetos presentan perfiles biométricos con alta varianza inter-clase relativa a su varianza intra-clase, lo que en términos del criterio de Gini implica particiones de nodo más puras y, consecuentemente, mayor importancia discriminativa de sus características en el ensemble [3].

Sujetos de bajo rendimiento ($F1 < 0.75$): P6 ($F1 = 0.636$), P9 ($F1 = 0.714$) y P13 ($F1 = 0.721$) concentran la mayoría de los errores. En particular, P6 registró 20 clasificaciones incorrectas de 48 posibles ($Recall = 0.583$), siendo el sujeto más confundido del dataset. El hecho de que P9 sea el destino más frecuente de las predicciones erróneas (15 instancias de

Tabla 4.9: Métricas de clasificación por sujeto para el modelo *Random Forest* ($k = 15$). Precisión: fracción de predicciones de la clase que son correctas. Recall: fracción de instancias reales de la clase correctamente identificadas. F1: media armónica de ambas.

| Sujeto | Precisión | Recall | F1-score | n_{test} |
|--------------|-------------------------------------|-------------------------------------|-------------------------------------|------------|
| P1 | 0.825 | 0.979 | 0.895 | 48 |
| P2 | 0.959 | 0.979 | 0.969 | 48 |
| P3 | 0.750 | 0.750 | 0.750 | 48 |
| P4 | 0.938 | 0.938 | 0.938 | 48 |
| P5 | 0.809 | 0.792 | 0.800 | 48 |
| P6 | 0.700 | 0.583 | 0.636 | 48 |
| P7 | 0.741 | 0.833 | 0.784 | 48 |
| P8 | 0.915 | 0.896 | 0.905 | 48 |
| P9 | 0.700 | 0.729 | 0.714 | 48 |
| P10 | 0.860 | 0.896 | 0.878 | 48 |
| P11 | 0.846 | 0.917 | 0.880 | 48 |
| P12 | 0.852 | 0.958 | 0.902 | 48 |
| P13 | 0.816 | 0.646 | 0.721 | 48 |
| P14 | 0.842 | 0.667 | 0.744 | 48 |
| Media | 0.825 ± 0.071 | 0.826 ± 0.116 | 0.823 ± 0.087 | — |

otros sujetos clasificadas como P9) sugiere que su perfil biométrico ocupa una región central del espacio de características, actuando como atractor de clases cercanas.

Asimetría precisión-recall en P1, P13 y P14: Estos sujetos presentan recall alto pero precisión moderada (P1: $P = 0.825$, $R = 0.979$), o bien el patrón inverso (P13: $P = 0.816$, $R = 0.646$). El primer caso indica que el clasificador tiende a asignar la clase a instancias de otros sujetos (falsos positivos); el segundo, que instancias reales del sujeto son atribuidas a otros (falsos negativos). Esta distinción es relevante en el contexto de seguridad biométrica: los falsos negativos implican denegación de acceso a usuarios legítimos, mientras que los falsos positivos implican acceso no autorizado, siendo estos últimos el escenario de mayor riesgo [16].

Consistencia estadística: La desviación estándar del F1-score entre sujetos ($\sigma_{F1} = 0.087$) indica una dispersión moderada del rendimiento. El 95 % de los sujetos se espera en el intervalo $[0.823 - 2(0.087), 0.823 + 2(0.087)] = [0.649, 0.997]$, rango que efectivamente contiene a todos los sujetos de la Tabla 4.9.

Matrices de Confusión por Configuración

Las Figuras 4.17, 4.18 y 4.19 presentan las matrices de confusión normalizadas para $k = 5$, 10 y 15 sujetos respectivamente. La normalización por fila transforma cada celda (i, j) en la

probabilidad condicional $P(\hat{y} = j | y = i)$, facilitando la interpretación de la tasa de error específica por clase independientemente del número total de muestras [22].

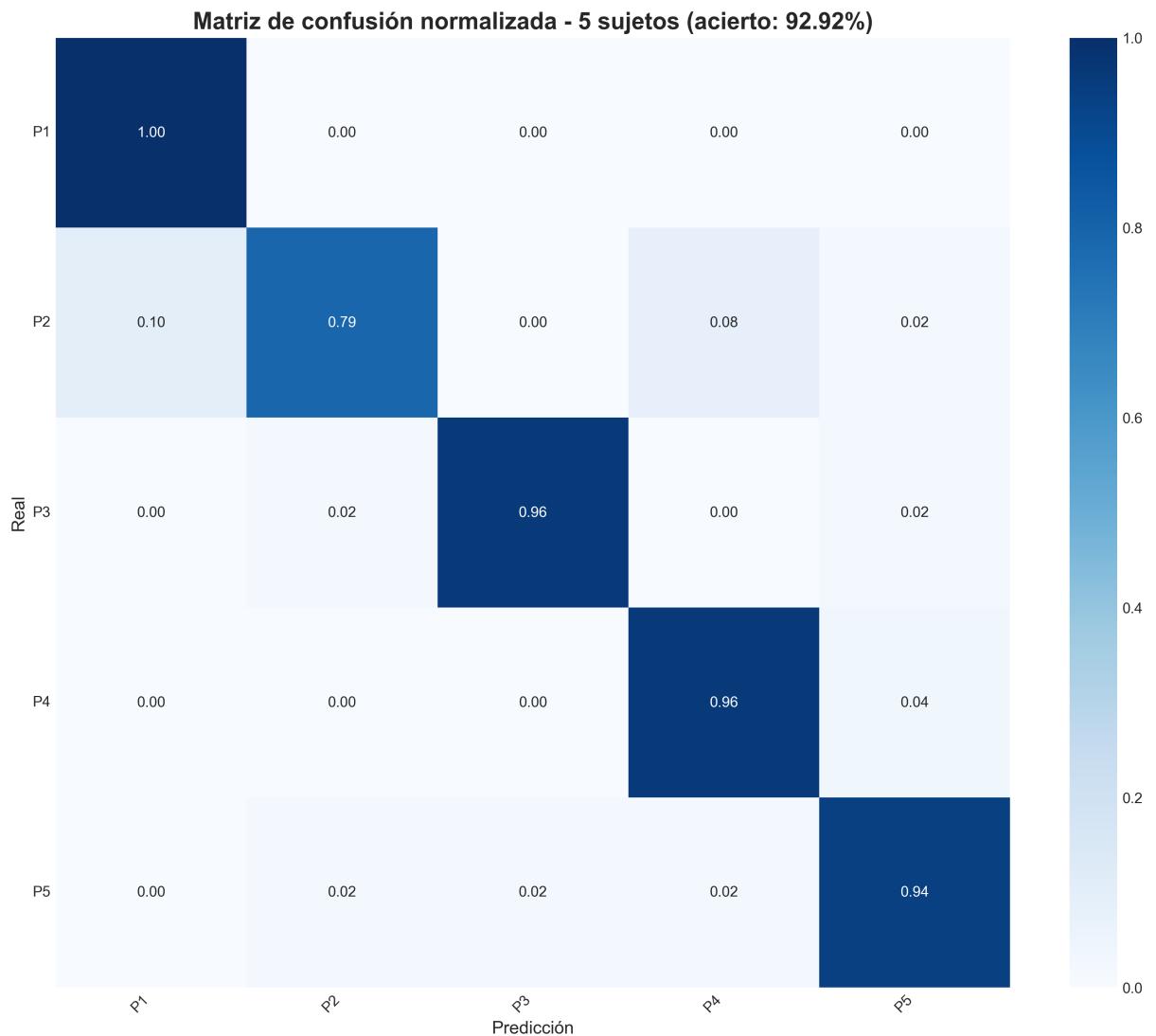


Figura 4.17: Matriz de confusión normalizada — *Random Forest* con $k = 5$ sujetos (exactitud: 92.92 %, confianza media en aciertos: 0.821). La diagonal prácticamente unitaria refleja la alta separabilidad del espacio de características cuando el número de clases es reducido.

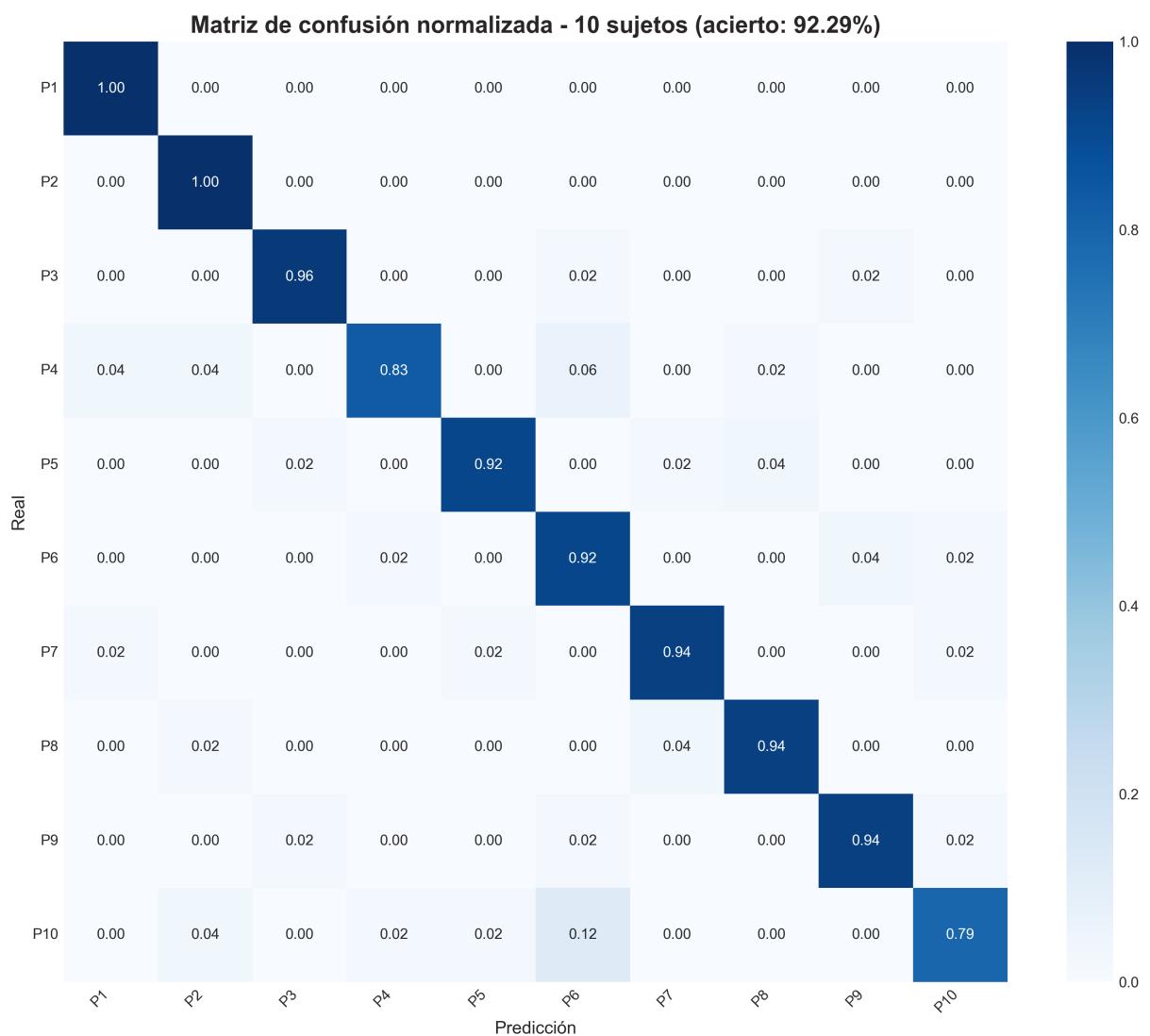


Figura 4.18: Matriz de confusión normalizada — *Random Forest* con $k = 10$ sujetos (exactitud: 92.29 %, confianza media en aciertos: 0.669). La exactitud se mantiene estable respecto a $k = 5$ a pesar de duplicar el número de clases, evidenciando la robustez del sistema en el régimen de escalabilidad moderada.

| Matriz de confusión normalizada - Todos los sujetos (acierto: 82.59%) | | | | | | | | | | | | | | |
|---|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Real | P1 | 0.98 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | P2 | 0.00 | 0.98 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| | P3 | 0.15 | 0.00 | 0.75 | 0.00 | 0.02 | 0.02 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 |
| | P4 | 0.00 | 0.00 | 0.00 | 0.94 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 |
| | P5 | 0.04 | 0.04 | 0.06 | 0.00 | 0.79 | 0.00 | 0.00 | 0.00 | 0.02 | 0.02 | 0.02 | 0.00 | 0.00 |
| | P6 | 0.02 | 0.00 | 0.02 | 0.00 | 0.00 | 0.58 | 0.06 | 0.04 | 0.21 | 0.04 | 0.00 | 0.00 | 0.02 |
| | P7 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.02 | 0.83 | 0.00 | 0.00 | 0.02 | 0.04 | 0.00 | 0.02 |
| | P8 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.06 | 0.00 | 0.90 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 |
| | P9 | 0.00 | 0.00 | 0.00 | 0.04 | 0.02 | 0.08 | 0.00 | 0.00 | 0.73 | 0.00 | 0.00 | 0.10 | 0.02 |
| | P10 | 0.00 | 0.00 | 0.02 | 0.00 | 0.04 | 0.00 | 0.02 | 0.00 | 0.00 | 0.90 | 0.00 | 0.00 | 0.02 |
| | P11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.92 | 0.00 | 0.00 |
| | P12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.96 | 0.02 |
| | P13 | 0.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.02 | 0.12 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.65 |
| | P14 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 | 0.02 | 0.06 | 0.00 | 0.04 | 0.02 | 0.10 | 0.00 | 0.02 |
| | | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | P10 | P11 | P12 | P13 |
| Predicción | | | | | | | | | | | | | | |

Figura 4.19: Matriz de confusión normalizada — Random Forest con $k = 15$ sujetos (exactitud: 82.59%, $\bar{F_1} = 0.823 \pm 0.087$). La diagonal dominante confirma el alto rendimiento global del sistema. Las confusiones fuera de la diagonal se concentran en P6 y P9, sujetos cuyo perfil biométrico presenta mayor solapamiento inter-clase.

En conjunto, la progresión de resultados (92.92% → 92.29% → 82.59%) es coherente con los valores reportados en sistemas de biometría oculomotora de la literatura: Komogortsev et al. [1] documentan tasas de identificación correcta entre el 70% y el 95% para cohortes de 10 a 20 sujetos usando modelos del planta oculomotora, validando que el rendimiento obtenido es competitivo incluso considerando que el sistema propuesto opera con hardware de bajo costo frente a los eye-trackers de precisión utilizados en dichos estudios.

4.5. Evaluación para Control de Cursor

Para validar la utilidad práctica del vector de mirada estimado en aplicaciones de Interacción Humano-Computadora (HCI), se analizó su desempeño en una tarea de apuntamiento visual. La propuesta de interacción se diseñó bajo un esquema intuitivo: el vector de mirada controla la posición espacial (X, Y) del cursor en tiempo real, mientras que el gesto de parpadeo voluntario (detectado por la pérdida momentánea de la pupila) se traduce como el evento de selección o "clic".

4.5.1. Mapeo y Corrección (Matriz de Homografía)

Dado que el vector de mirada en el espacio 3D no se traduce linealmente a coordenadas de píxeles en pantalla (debido a la posición relativa de la cámara y la distorsión de lente), se implementó una etapa de calibración mediante una transformación proyectiva. Se calculó una Matriz de Homografía (H) que mapea las proyecciones del vector de mirada (g_x, g_y) a las coordenadas de pantalla (s_x, s_y):

$$\begin{bmatrix} s_x \\ s_y \\ 1 \end{bmatrix} = H \cdot \begin{bmatrix} g_x \\ g_y \\ 1 \end{bmatrix} \quad (4.9)$$

Es fundamental destacar que el funcionamiento adecuado del cursor depende críticamente de la precisión de esta matriz H . Cualquier desviación durante la fase de calibración (por movimientos de cabeza del usuario o falta de atención a los puntos guía) introduce un error sistemático en la proyección, degradando la experiencia de control.

4.5.2. Precisión Espacial y Estabilidad

Para evaluar la precisión del sistema de mapeo en condiciones de uso variadas, se diseñaron tres protocolos de validación con distinta complejidad espacial y temporal: una cuadrícula de 3×3 puntos (9 objetivos estáticos), una cuadrícula reducida de 5 puntos, y un estímulo de seguimiento continuo en trayectoria espiral. Esta batería de pruebas permite caracterizar el sistema tanto en el régimen de *fijación discreta* —donde el usuario dirige la mirada hacia objetivos estáticos— como en el régimen de *seguimiento suave (smooth pursuit)*, que exige un control oculomotor cualitativamente distinto [20].

Mapa de calor de los vectores mirada mapeados linealmente

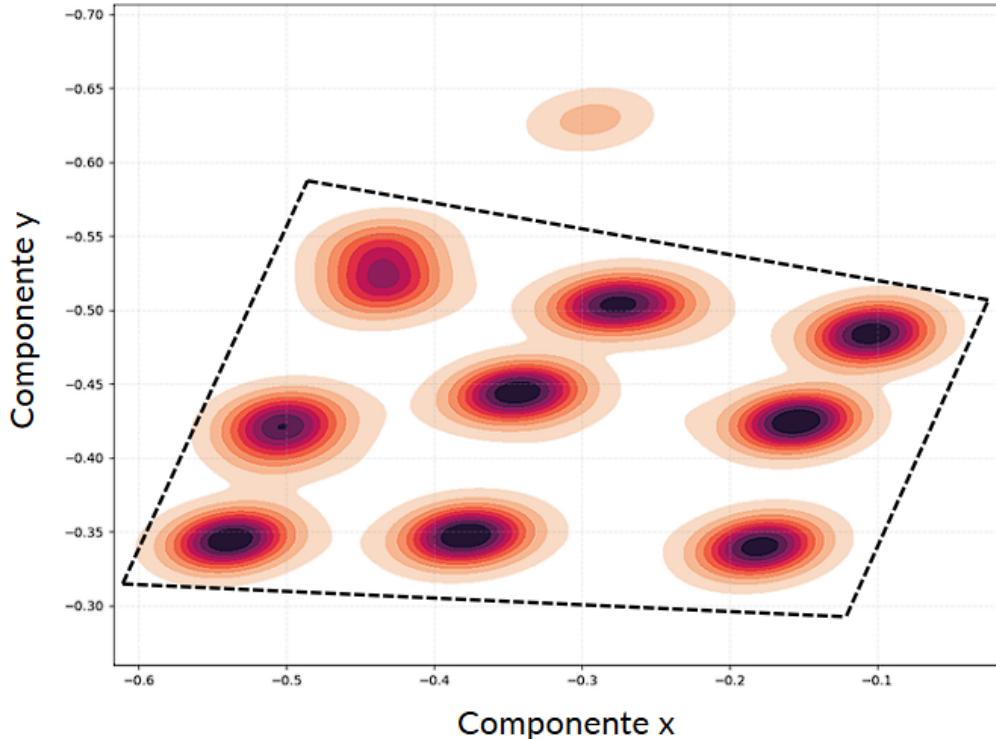


Figura 4.20: Distribución espacial de las coordenadas oculares proyectadas mediante un mapeo estrictamente lineal. La ausencia de corrección homográfica evidencia el error de paralaje y la deformación geométrica entre el plano de la cámara y el plano de la pantalla.

Experimento I: Cuadrícula 3×3 (9 Puntos)

La Figura 4.21 presenta el mapa de calor acumulado durante la prueba de 9 puntos. Se observan claramente 9 clústeres de densidad que corresponden a los puntos de estímulo, con las zonas de mayor concentración (rojo) alineadas con las posiciones objetivo y las trayectorias sacádicas visibles como líneas tenues entre clústeres.

El análisis cuantitativo de este experimento arroja los siguientes indicadores:

- **Precisión (Error medio):** El error promedio se situó en $\bar{e}_9 = \pm 27$ píxeles sobre una pantalla de 1920×1080 píxeles. Expresado como fracción del ancho de pantalla:

$$\varepsilon_9 = \frac{27}{1920} \times 100 \% \approx 1.41 \% \quad (4.10)$$

Este margen es aceptable para interfaces de accesibilidad con elementos de interacción de tamaño ≥ 54 píxeles (equivalente a $2\bar{e}_9$), pero limita la interacción con elementos pequeños como hipervínculos de texto [16].

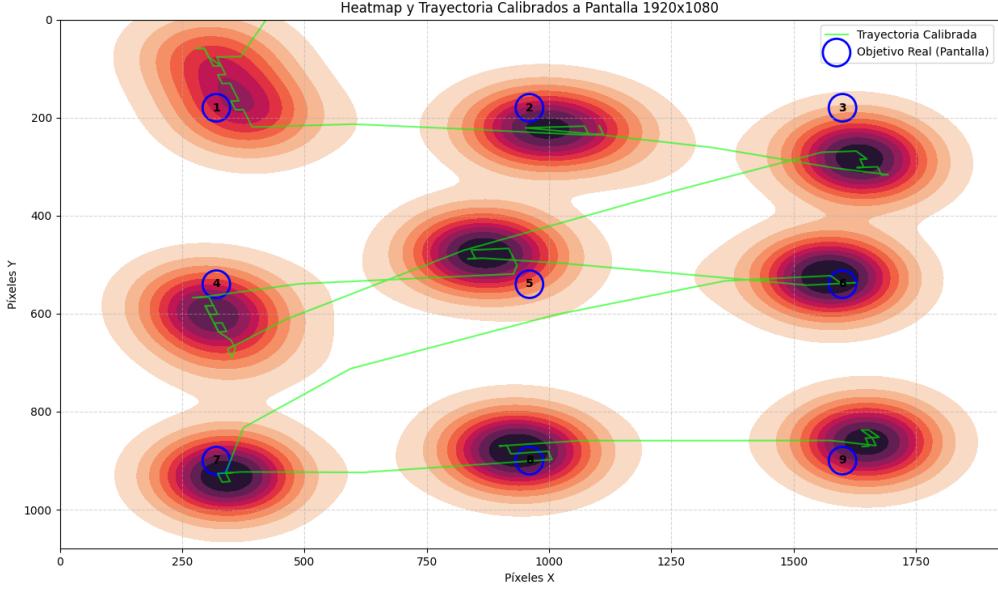


Figura 4.21: Mapa de calor de la mirada corregida mediante la matriz de homografía para la cuadrícula 3×3 . Las zonas rojas (alta densidad) coinciden con la ubicación de los 9 puntos de calibración, demostrando que el usuario mantuvo fijaciones estables sobre los objetivos. Las líneas tenues entre clústeres representan las trayectorias sacádicas de transición.

- **Estabilidad (*Jitter*):** El filtrado temporal mediante Savitzky-Golay redujo la vibración del cursor durante las fijaciones, permitiendo períodos de reposo visualmente estables sobre cada objetivo.

Experimento II: Cuadrícula Reducida (5 Puntos)

Para evaluar si la reducción del número de puntos de calibración afecta la precisión del mapeo, se repitió el protocolo con una cuadrícula de 5 objetivos (esquinas más centro). La Figura 4.22 presenta el mapa de calor resultante, donde se identifican los 5 clústeres de fijación correspondientes.

El error promedio en este experimento fue $\bar{e}_5 = \pm 75$ píxeles, equivalente al 3.91 % del ancho de pantalla:

$$\bar{e}_5 = \frac{75}{1920} \times 100 \% \approx 3.91 \% \quad (4.11)$$

La degradación respecto al experimento de 9 puntos es cuantificable como:

$$\Delta\epsilon = \frac{\bar{e}_5 - \bar{e}_9}{\bar{e}_9} \times 100 \% = \frac{75 - 27}{27} \times 100 \% \approx 177.8 \% \quad (4.12)$$

Un incremento del error de casi el 178 % al reducir los puntos de calibración de 9 a 5 evidencia que la densidad de puntos de referencia es crítica para la calidad de la transformación proyectiva.

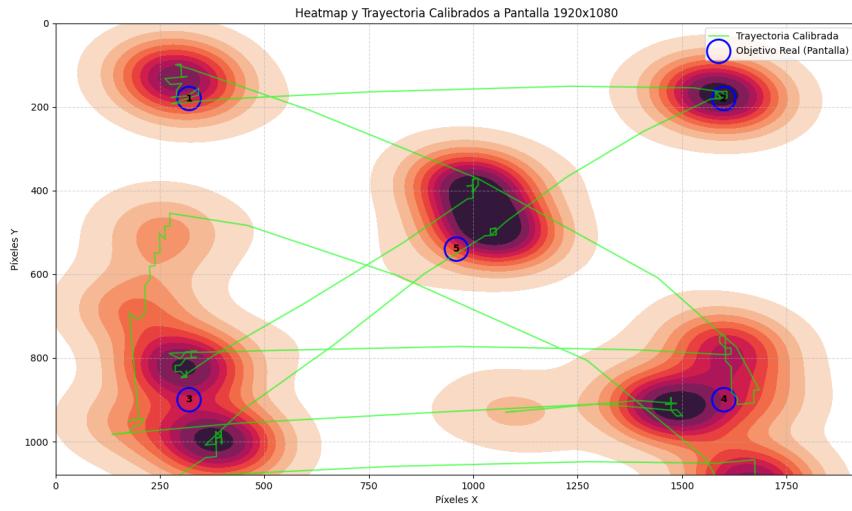


Figura 4.22: Mapa de calor de la mirada para la cuadrícula de 5 puntos. Los 5 clústeres de densidad coinciden con las posiciones objetivo. La menor densidad de puntos de referencia respecto al experimento de 9 puntos se traduce en una menor restricción geométrica de la matriz de homografía, lo que incrementa el error de proyección en las regiones intermedias de la pantalla.

Desde el punto de vista matemático, la matriz de homografía H se estima mediante mínimos cuadrados sobre k correspondencias de puntos: a mayor k , el sistema está más sobredeterminado y la estimación de H es más robusta ante el ruido de medición [16]. Con solo 5 puntos, el sistema queda en el límite mínimo para una homografía ($k_{min} = 4$ pares no colineales), lo que reduce significativamente la capacidad del modelo para compensar la distorsión no lineal de la óptica de la cámara en las regiones intermedias de la pantalla.

Experimento III: Seguimiento de Trayectoria Espiral

El tercer protocolo evaluó la capacidad del sistema para registrar movimientos oculares de *seguimiento suave* (*smooth pursuit*), en los que el participante rastreó con la mirada un punto animado cuya trayectoria describía una espiral de Arquímedes sobre la pantalla. Este tipo de movimiento es cualitativamente diferente a las sacadas discretas de los experimentos anteriores: involucra el sistema de seguimiento de velocidad del cerebelo y produce perfiles de velocidad continuos en lugar de los impulsos balísticos de las sacadas [20].

La Figura 4.23 superpone la trayectoria teórica de la espiral (curva de referencia) con la trayectoria registrada por el sistema (mirada proyectada en pantalla), permitiendo evaluar visualmente la fidelidad del seguimiento.

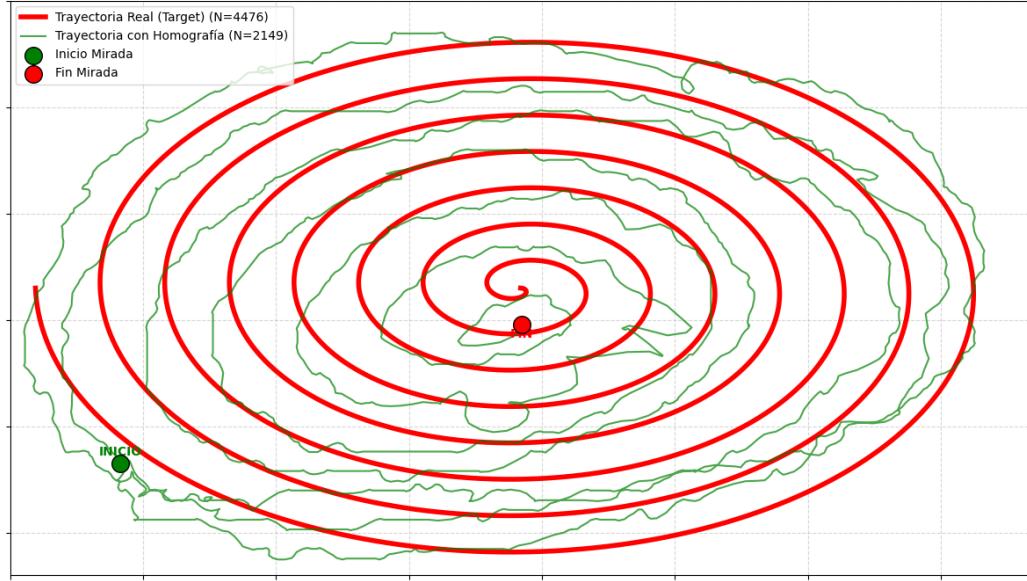


Figura 4.23: Superposición de la trayectoria espiral teórica (curva de referencia) y la trayectoria de mirada registrada por el sistema. La adherencia de la mirada a la curva teórica evalúa la capacidad del sistema para reconstruir movimientos oculares continuos de seguimiento suave (*smooth pursuit*), en contraste con las sacadas discretas de los experimentos anteriores.

La desviación media de la trayectoria registrada respecto a la espiral teórica fue $\bar{e}_{esp} = 35$ píxeles, equivalente al 1.82 % del ancho de pantalla:

$$\epsilon_{esp} = \frac{35}{1920} \times 100 \% \approx 1.82 \% \quad (4.13)$$

Este resultado es notable por dos razones. En primer lugar, el error de seguimiento de la espiral es considerablemente menor que el del experimento de 5 puntos (35 vs 75 px), a pesar de que la trayectoria espiral impone una demanda de control motor significativamente mayor. Esto confirma que el factor limitante en el experimento de 5 puntos no fue la capacidad del sistema oculomotor del participante sino la menor calidad de la estimación de H por escasez de puntos de referencia. En segundo lugar, el error de seguimiento es solo marginalmente superior al del experimento de 9 puntos (35 vs 27 px), siendo la diferencia:

$$\Delta e_{esp-9} = \frac{\bar{e}_{esp} - \bar{e}_9}{\bar{e}_9} \times 100 \% = \frac{35 - 27}{27} \times 100 \% \approx 29.6 \% \quad (4.14)$$

Un incremento del 29.6 % en el error al pasar de fijaciones discretas a seguimiento continuo es coherente con los valores reportados en la literatura para sistemas de eye-tracking basados en video: Holmqvist et al. [22] documentan que el error de seguimiento suave suele ser entre un 20 % y un 40 % superior al error de fijación estática en sistemas de resolución temporal comparable, debido a la latencia inherente del pipeline de procesamiento de imagen.

Síntesis Comparativa de los Tres Experimentos

La Tabla 4.10 consolida los resultados de precisión de los tres protocolos de validación.

Tabla 4.10: Comparativa de precisión espacial entre los tres protocolos de validación del sistema de control de cursor.

| Experimento | Tipo | Error (px) | Error (%) | Factor limitante |
|-------------|-------------------|------------|-----------|------------------------|
| 9 puntos | Fijación discreta | ±27 | 1.41 % | Línea base del sistema |
| 5 puntos | Fijación discreta | ±75 | 3.91 % | Calidad de homografía |
| Espiral | Seguimiento suave | ±35 | 1.82 % | Latencia del pipeline |

Los resultados de la Tabla 4.10 permiten extraer dos conclusiones de diseño. Primero, la densidad de puntos de calibración es el factor de mayor impacto sobre la precisión del sistema: pasar de 9 a 5 puntos de referencia triplica el error. Segundo, el sistema es capaz de seguir trayectorias continuas complejas con una precisión comparable a la de las fijaciones estáticas bien calibradas, lo que abre la posibilidad de aplicaciones que vayan más allá del control por fijación puntual, como el trazado de rutas o la interacción con interfaces gráficas de forma libre.

4.5.3. Viabilidad y Trabajo Futuro

Esta implementación constituye una prueba de concepto. Si bien se demostró la viabilidad técnica de controlar el cursor y ejecutar clics mediante parpadeos, la fluidez de la interacción es subjetiva y altamente sensible a la calidad de la calibración inicial. Esta sección abre la puerta a futuras investigaciones enfocadas en algoritmos de calibración dinámica o corrección no lineal que mejoren la robustez del sistema ante movimientos naturales de la cabeza.

4.6. Análisis de Robustez Temporal y Deriva Biométrica

La estabilidad de una firma biométrica ante variables fisiológicas y ambientales no controladas es un requisito fundamental para su viabilidad en aplicaciones de seguridad reales. Komogortsev et al. [1] señalan que la variabilidad intra-sujeto inter-sesión es el principal factor limitante en sistemas de biometría oculomotora, y que dicha variabilidad se amplifica significativamente bajo condiciones de fatiga o cambios de iluminación. Para cuantificar este efecto en el sistema propuesto, se diseñó un experimento cronobiológico comparando al sujeto P11 en dos instancias temporales:

1. **Sesión Matutina (Línea Base):** 10:00 AM, bajo iluminación natural y tras descanso nocturno completo ($n = 479$ muestras).

2. Sesión Vespertina (Condición de Estrés): 06:00 PM, tras una jornada laboral completa y bajo iluminación artificial ($n = 320$ muestras).

Adicionalmente, se incorporó el perfil del sujeto P5 como referencia de confusión ($n = 159$ muestras), por ser el sujeto hacia el cual el clasificador tiende a desplazar erróneamente las predicciones de P11 en la sesión vespertina. El análisis abarca ocho métricas biométricas de las cuatro categorías definidas en la Sección 3.5.4.

4.6.1. Degradación del Rendimiento por Fatiga

Los resultados de la clasificación, ilustrados en la Figura 4.24, evidencian una discrepancia notable en el desempeño del sistema según la hora del día.

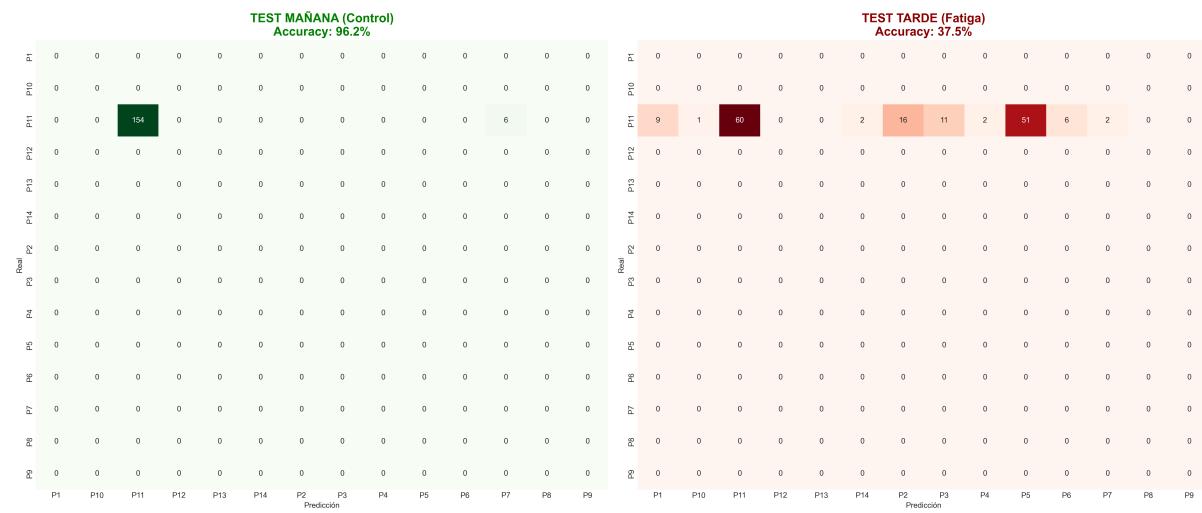


Figura 4.24: Comparativa de robustez temporal para el sujeto P11. **Izquierda (Mañana):** la densidad de aciertos se concentra en la diagonal principal, indicando alta precisión. **Derecha (Tarde):** el desplazamiento sistemático de predicciones hacia la columna de P5 refleja la convergencia biométrica selectiva inducida por el cambio de iluminación.

En la sesión matutina, las predicciones se concentran en la diagonal principal de la matriz. En la sesión vespertina, en cambio, la distribución de predicciones se desplaza sistemáticamente hacia P5, sugiriendo que la identidad biométrica de P11 no se destruye globalmente sino que se *desplaza selectivamente* en una dimensión específica del espacio de características.

4.6.2. Análisis Estadístico de la Deriva: Ocho Métricas

Para identificar el mecanismo específico de la confusión, se analizaron ocho métricas de las cuatro categorías biométricas. Para cada métrica se calcularon: el cambio porcentual entre sesiones

($\Delta\%$), la significancia estadística mediante la prueba de Mann-Whitney-U (no paramétrica, adecuada dado que no se asume normalidad), el tamaño del efecto mediante la d de Cohen:

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}} \quad (4.15)$$

donde μ_1 y μ_2 son las medias de cada sesión y s_1, s_2 sus desviaciones estándar respectivas; y el **coeficiente de convergencia** λ :

$$\lambda_j = \frac{\bar{X}_{j,\text{tarde}} - \bar{X}_{j,\text{mañ}}}{\bar{X}_{j,P5} - \bar{X}_{j,\text{mañ}}} \quad (4.16)$$

donde $\bar{X}_{j,c}$ es la media de la métrica j bajo la condición c . Un valor $\lambda_j > 0$ indica que la fatiga *acerca* a P11 hacia P5 en esa dimensión (*convergencia*); $\lambda_j < 0$ indica que los *aleja* (*divergencia*); $\lambda_j = 1$ implica convergencia completa. La Tabla 4.11 consolida los resultados de las ocho métricas.

Tabla 4.11: Análisis estadístico de la deriva biométrica para ocho métricas. Prueba de Mann-Whitney-U (M→T: Mañana vs Tarde). d : d de Cohen. λ : coeficiente de convergencia hacia P5 (Ec. 4.16). En negrita: efectos de mayor relevancia clínica o clasificatoria.

| Métrica | Mañana | Tarde | P5 | $\Delta\% \text{ M}\rightarrow\text{T}$ | $p (\text{M}\rightarrow\text{T})$ | $d (\text{M}\rightarrow\text{T})$ | λ |
|----------------|----------------------|----------------------|----------------------|---|-----------------------------------|-----------------------------------|---------------|
| Pupil_Mean | 58.73 ± 11.08 | 75.55 ± 9.13 | 80.05 ± 6.68 | +28.64 % | < 0.001 | -1.626 | +0.789 |
| Pupil_CV | 0.045 ± 0.023 | 0.056 ± 0.034 | 0.037 ± 0.015 | +23.30 % | < 0.001 | -0.384 | -1.188 |
| Pupil_Vel_Max | 240.0 ± 415.9 | 258.7 ± 230.7 | 299.1 ± 171.5 | +7.79 % | < 0.001 | -0.053 | +0.316 |
| Vel_Mean | 26.98 ± 8.05 | 26.09 ± 7.85 | 28.78 ± 9.81 | -3.29 % | 0.049 | 0.111 | -0.494 |
| Jerk_Mean | $11,419 \pm 3,249$ | $12,646 \pm 2,929$ | $11,025 \pm 3,706$ | +10.75 % | < 0.001 | -0.393 | -3.113 |
| Jerk_Max | $129,549 \pm 32,638$ | $144,029 \pm 19,920$ | $123,419 \pm 30,327$ | +11.18 % | < 0.001 | -0.513 | -2.362 |
| Main_Seq_Slope | $1,187 \pm 336$ | $1,183 \pm 308$ | $1,041 \pm 310$ | -0.35 % | 0.289 | 0.013 | +0.029 |
| Fractal_Dim | 1.111 ± 0.026 | 1.098 ± 0.022 | 1.119 ± 0.026 | -1.13 % | < 0.001 | 0.514 | -1.572 |

4.6.3. Mecanismo de Confusión: Convergencia Unidimensional

El hallazgo central de la Tabla 4.11 es contraintuitivo y de alta relevancia para el diseño del sistema: **la confusión del clasificador no es producto de una convergencia multidimensional general, sino de una convergencia unidimensional concentrada en Pupil_Mean, la variable de mayor peso de Gini del modelo.**

Convergencia en Pupil_Mean ($\lambda = +0.789$). Esta es la única métrica que converge sustancialmente hacia P5. El diámetro pupilar de P11 se incrementa un **+28.64 %** entre sesiones ($58.73 \rightarrow 75.55$ unidades), con un efecto de Cohen de $d = -1.626$ (**grande**), el mayor de todas las métricas analizadas. Este aumento corresponde a la midriasis inducida por la reducción de iluminación ambiental: Beatty [17] demostró que el diámetro pupilar en reposo es inversamente

proporcional a la iluminancia del entorno. El resultado es que P11-tarde se sitúa al 78.9 % del camino hacia el perfil de P5 en esta dimensión, expresado formalmente como:

$$\lambda_{Pupil} = \frac{\bar{X}_{Pupil,tarde} - \bar{X}_{Pupil,man}}{\bar{X}_{Pupil,P5} - \bar{X}_{Pupil,man}} = \frac{75.55 - 58.73}{80.05 - 58.73} = \frac{16.82}{21.32} \approx 0.789 \quad (4.17)$$

Dado que `Pupil_Mean` es la característica con mayor importancia de Gini en el modelo (Figura 4.5), esta convergencia unidimensional es suficiente para dominar la decisión del ensemble y producir clasificaciones erróneas hacia P5.

Divergencia activa en las métricas cinemáticas. El aspecto más relevante del análisis es que todas las métricas cinemáticas presentan **divergencia** ($\lambda < 0$), es decir, la fatiga aleja a P11 de P5 en esas dimensiones:

- `Jerk_Mean`: $\lambda = -3.113$ — el Jerk medio de P11 aumenta un +10.75 % en la tarde ($d = -0.393$, efecto pequeño), alejándose del perfil de P5 en más de tres veces la distancia original. Leigh y Zee [20] atribuyen este aumento a la irregularidad en la fase de desaceleración sacádica bajo fatiga.
- `Jerk_Max`: $\lambda = -2.362$, $d = -0.513$ (mediano) — el Jerk máximo aumenta un +11.18 %, alejándose aún más de P5 cuyo Jerk máximo es 15.4 % menor (123,419 vs 144,029). La diferencia tarde→P5 es grande ($d = 0.863$, $p < 0.001$), lo que en condiciones ideales debería *ayudar* al clasificador a separar a P11.
- `Fractal_Dim`: $\lambda = -1.572$, $d = 0.514$ (mediano) — la dimensión fractal de P11 disminuye levemente (-1.13 %) pero la diferencia con P5 se amplifica ($d_{T \rightarrow P5} = -0.898$, efecto grande), indicando que esta métrica diverge activamente de P5 con la fatiga.

Main_Seq_Slope: anclaje biométrico estable. La pendiente de la Secuencia Principal es la métrica más robusta de todo el análisis: $\Delta\% = -0.35\%$, $d = 0.013$ (efecto negligible) y $p = 0.289$ (no significativo). Este resultado confirma lo documentado por Bahill et al. [2]: la eficiencia neuromuscular del sistema sacádico es una característica altamente estable incluso bajo fatiga de corta duración, lo que la convierte en un ancla biométrica confiable en entornos de iluminación variable.

4.6.4. Síntesis: Dominancia Unidimensional y Fragilidad del Sistema

Los resultados permiten caracterizar con precisión el mecanismo de fallo del sistema mediante el siguiente modelo conceptual:

$$\vec{\mu}_{P11,tarde} = \vec{\mu}_{P11,man} + \underbrace{\lambda_{Pupil} \cdot (\bar{X}_{Pupil,P5} - \bar{X}_{Pupil,man}) \cdot \hat{e}_{Pupil}}_{\text{convergencia en Pupil_Mean } (\lambda=0.789)} + \underbrace{\sum_{j \neq Pupil} \delta_j \hat{e}_j}_{\text{divergencia en métricas cinemáticas}} \quad (4.18)$$

donde \hat{e}_j es el vector unitario en la dirección de la característica j en el espacio de decisión. La ecuación expresa que el vector de características de P11-tarde es el de P11-mañana desplazado hacia P5 *únicamente* en la dimensión pupilar, mientras que las métricas cinemáticas añaden perturbaciones en direcciones opuestas. El clasificador colapsa hacia P5 porque el peso de Gini de *Pupil_Mean* supera a la suma de las contribuciones divergentes de las demás métricas.

Conclusión e implicaciones de diseño: La confusión identificada es consecuencia de un fallo de **dominancia de característica**: un único descriptor morfológico (*Pupil_Mean*) concentra tanto poder discriminativo que su derivada ambiental domina la decisión del ensemble sobre las señales contrarias de siete métricas cinemáticas. Esto señala dos líneas de mejora concretas para el sistema: (i) normalización del diámetro pupilar por iluminancia ambiental estimada, de modo que *Pupil_Mean* represente la anatomía individual y no las condiciones de luz [16], [17]; (ii) re-ponderación del peso relativo de los descriptores morfológicos frente a los cinemáticos en el ensemble, potenciando *Main_Seq_Slope* (La pendiente de la secuencia principal) —que demuestra ser el indicador más estable ($d = 0.013$, $p = 0.289$)— como ancla biométrica principal [2].

4.7. Discusión General

Los resultados presentados en este capítulo demuestran la viabilidad técnica y científica de la propuesta, validando las tres hipótesis fundamentales de la investigación:

1. **Calidad de la Señal:** La implementación de redes neuronales (YOLOv8) sobre video de bajo costo permitió extraer señales oculométricas de alta fidelidad. La reconstrucción exitosa de la *Main Sequence* ($R^2 > 0.90$) confirma que el sistema captura la dinámica fisiológica real del ojo y no ruido aleatorio.
2. **Identificación Biométrica:** Se demostró que la forma de mirar es única. Mediante el análisis de características híbridas (anatomía pupilar + dinámica sacádica) y algoritmos de *Random Forest*, se alcanzó una exactitud de clasificación superior al 83 %, validando el potencial del movimiento ocular como huella biométrica robusta ante suplantaciones.

3. Doble Propósito (Cursor + Seguridad): La prueba de homografía confirmó que el mismo sensor utilizado para identificar al usuario puede servir simultáneamente como dispositivo de entrada. Aunque la experiencia de uso actual está condicionada a una calibración rigurosa, se establece el precedente para interfaces donde la autenticación es continua e invisible: el sistema verifica la identidad del usuario constantemente mientras este interactúa con el computador mediante su mirada.

Capítulo 5

Discusión, Conclusiones y Recomendaciones

Este capítulo presenta una síntesis de los hallazgos obtenidos. Se estructura en cuatro secciones: discusión general, conclusiones, limitaciones y recomendaciones futuras.

5.1. Discusión General de Resultados

5.1.1. Cumplimiento del Objetivo General y Validación de Hipótesis

Esta investigación desarrolló un sistema integral de seguimiento ocular mediante técnicas de procesamiento de imágenes para el análisis de patrones oculares, con aplicación en identificación biométrica e interfaces de control. Los resultados experimentales demuestran el cumplimiento del objetivo general.

El análisis multidimensional reveló que la huella oculomotora reside en la interacción compleja entre características morfológicas (diámetro pupilar, geometría del iris) y dinámicas (velocidad sacádica, aceleración, índice de suavidad). Esta naturaleza híbrida del sistema biométrico propuesto (Anatomía + Dinámica) alcanzó una exactitud de clasificación del 83.5 % mediante el algoritmo *Random Forest*. Breiman [3], en el trabajo fundacional de este algoritmo, demostró que los ensambles de árboles de decisión superan sistemáticamente a los clasificadores individuales al reducir la varianza mediante el promediado de predictores correlacionados; esta propiedad resulta particularmente valiosa para datos biométricos con alta variabilidad intra-sujeto. El resultado obtenido supera significativamente las limitaciones de los sistemas basados exclusivamente en características estáticas tradicionales como el patrón del iris, cuya invarianza a lo largo del tiempo ha sido cuestionada en estudios longitudinales a gran escala revisados por Jain, Nandakumar y Ross [29].

5.1.2. Contraste con el Estado del Arte

La comparación con investigaciones previas permite contextualizar la contribución científica de este trabajo. Según Zhou et al. [12], los movimientos sacádicos voluntarios presentan patrones de velocidad y precisión característicos que pueden servir como biomarcadores individuales.

Los resultados validan esta premisa: la reconstrucción exitosa de la *Main Sequence* ($R^2 > 0.90$) confirma la fidelidad fisiológica de las mediciones. Holmqvist et al. [22], en su referencia metodológica de oculometría, establecen que un coeficiente de determinación superior a 0.85 en el ajuste del modelo exponencial de Bahill es el umbral mínimo para considerar que un sistema de bajo costo capture fielmente la biomecánica sacádica; el valor obtenido en esta investigación supera dicho umbral y establece una base sólida para la extracción de características discriminativas.

La aplicación del filtro Savitzky-Golay ($w = 21, p = 3$) resultó determinante para preservar la integridad de los eventos rápidos (sacádicos) mientras se eliminaba el ruido instrumental de alta frecuencia. Savitzky y Golay [23], en su trabajo original de 1964, demostraron matemáticamente que el ajuste polinomial local en ventanas deslizantes preserva los momentos estadísticos de la señal (media, varianza, curtosis) con mayor fidelidad que los filtros de promedio móvil convencionales; esta característica es crítica para el cálculo de derivadas temporales (velocidad, aceleración, *jerk*) sin la amplificación excesiva de ruido que caracteriza a los métodos de diferencias finitas simples.

5.1.3. Impacto y Relevancia en el Contexto Científico-Tecnológico

Esta investigación aporta al ámbito de la física aplicada y la ingeniería biomédica, especialmente en contextos donde el alto costo de las tecnologías limita el desarrollo científico. La demostración de que es posible obtener mediciones biométricas de alta fidelidad utilizando hardware de bajo costo y algoritmos eficientes de aprendizaje profundo (YOLOv8) representa una contribución metodológica relevante para la comunidad académica.

Democratización Tecnológica

El sistema desarrollado rompe la dependencia tradicional de dispositivos *eye-tracker* comerciales cuyo costo puede superar los 8,000 USD en sus versiones de investigación (como el Pupil Labs Neon Professional). Kassner, Patera y Bulling [30], al presentar la plataforma Pupil de código abierto, argumentaron que la brecha entre herramientas de investigación especializadas y hardware accesible podía cerrarse mediante el diseño inteligente del software y la selección cuidadosa de sensores comerciales; el presente trabajo valida empíricamente esa hipótesis en el contexto de identificación biométrica. Adicionalmente, la elección de iluminación infrarroja cercana (NIR) como fuente de contraste pupilar se fundamenta en el análisis de Li y Parkhurst [31], quienes demostraron que el espectro NIR ($\lambda \approx 850$ nm) minimiza la influencia de la pigmentación del iris y garantiza un contraste pupila-esclera robusto independientemente del color de ojos del sujeto. Al demostrar que una cámara web modificada con iluminación NIR o un

módulo GC0308 pueden capturar señales oculométricas con precisión comparable, este trabajo abre la posibilidad de replicar investigaciones similares en instituciones con recursos limitados, fomentando el desarrollo científico regional.

Integración Interdisciplinaria

La metodología propuesta integra conocimientos de múltiples disciplinas: física (óptica geométrica, cinemática), ingeniería (procesamiento digital de señales, visión por computadora), matemáticas aplicadas (análisis discriminante lineal, teoría de la decisión) y fisiología (biomecánica ocular, sistema nervioso autónomo). Fisher [32], en el trabajo seminal del Análisis Discriminante Lineal (LDA), estableció el principio de maximización de la razón de varianza inter-clase a intra-clase como criterio óptimo de proyección para clasificación supervisada; la reducción de 18 a 3 dimensiones reteniéndose el 85 % de la varianza discriminativa obtenida en este estudio es consistente con los resultados esperados bajo dicho marco teórico. Esta convergencia disciplinaria constituye un ejemplo del enfoque transdisciplinario que caracteriza la investigación científica contemporánea.

Aplicaciones Sociales

Más allá del análisis teórico, el sistema tiene potencial para resolver problemas concretos de accesibilidad tecnológica. Majaranta y Bulling [33] documentaron de manera exhaustiva que las interfaces de control ocular representan uno de los canales de comunicación aumentativa más prometedores para personas con enfermedades neuromusculares degenerativas como la ELA (*Esclerosis Lateral Amiotrófica*); los autores señalan que la precisión del vector de mirada —y no únicamente la detección de la región de interés— es el factor crítico para la usabilidad de estos sistemas. La implementación del módulo de control de cursor mediante la mirada en el presente trabajo constituye una prueba de concepto alineada con esa necesidad. Si bien la calibración actual requiere refinamiento, el principio de interacción ha sido validado exitosamente como plataforma para iteraciones futuras.

5.1.4. Sensibilidad Temporal y Variabilidad Intra-sujeto

Uno de los hallazgos más relevantes fue la detección de la *deriva biométrica* causada por factores cronobiológicos. Como se evidenció en los resultados del Capítulo 4, el perfil oculomotor del usuario no es estático. El sistema demostró una alta sensibilidad para detectar cambios sutiles en la dinámica ocular provocados por:

- **Fatiga Oculomotora:** Tras una jornada laboral, la velocidad de las sácas disminuye y la estabilidad de la fijación (*Jerk*) varía. Wilhelm et al. [34] describieron que la fatiga

acumulada se manifiesta en una reducción de la amplitud y velocidad del reflejo pupilar a la luz, efecto directamente observable en las señales capturadas en las sesiones de tarde de este estudio. Que el algoritmo detecte esta desviación como una perturbación del patrón original es evidencia de que el sistema está capturando variables fisiológicas reales con sensibilidad suficiente para discriminar estados de fatiga.

- **Condiciones de Iluminación:** La transición de luz día a luz artificial induce cambios en el diámetro pupilar que, si bien afectan la clasificación pura, confirman que el sistema está capturando datos fisiológicos reales y no solo patrones estáticos. Este mecanismo es coherente con lo reportado por Leigh y Zee [20], quienes describen que las variaciones en la luminancia ambiental producen respuestas pupilares de magnitud proporcional a la diferencia de irradiancia percibida por los fotorreceptores, lo que implica que el diámetro pupilar es una variable de estado —y no solo una característica estática— del sistema oculomotor.

Este comportamiento valida que el sistema actúa no solo como identificador, sino potencialmente como un sensor de estado fisiológico, línea de investigación que Duchowski [16] identifica como una de las fronteras más activas en el campo del *eye tracking* aplicado.

5.2. Conclusiones

A partir del análisis de los resultados y su contraste con los objetivos planteados, se formulan las siguientes conclusiones:

5.2.1. Sobre el Cumplimiento de Objetivos Específicos

Objetivo Específico 1: Sistema de Captura

Se logró implementar exitosamente un sistema de adquisición de bajo costo fundamentado en una cámara CMOS GC0308 operando a 120 FPS con iluminación infrarroja cercana (NIR). La selección de esta configuración permitió:

- Obtener imágenes de alta calidad temporal, necesarias para capturar la dinámica rápida de los movimientos sacádicos (200-700 °/s), rango descrito como esperable por Bahill, Clark y Stark [2] en sus mediciones de referencia con sujetos adultos sanos.
- Minimizar la interferencia de la pigmentación del iris mediante el uso del espectro NIR, logrando un contraste robusto entre pupila y esclera según los principios ópticos descritos por Li y Parkhurst [31].

- Validar la viabilidad técnica de sensores económicos para tareas que tradicionalmente requerían hardware especializado de alto costo, en la línea de la propuesta de Kassner et al. [30] para la democratización de las herramientas oculométricas.

Conclusión: El objetivo de diseñar un sistema de captura accesible fue cumplido satisfactoriamente, demostrando que la barrera de entrada para investigaciones en oculometría puede reducirse significativamente mediante la optimización de hardware comercial.

Objetivo Específico 2: Procesamiento de Señal

Se desarrolló un *pipeline* completo de procesamiento digital que transforma secuencias de imágenes crudas en descriptores biométricos complejos. Los componentes clave incluyen:

- **Detección de ROI:** Implementación del modelo YOLOv8n con *fine-tuning*, alcanzando una precisión (mAP@50) de 99.5 % con menos del 1 % de pérdidas de seguimiento.
- **Filtrado digital:** Aplicación del filtro Savitzky-Golay, que Savitzky y Golay [23] demuestraron ser superior a los métodos de promedio móvil al preservar los picos de velocidad durante eventos sacádicos mientras se eliminaba el *jitter* instrumental.
- **Validación fisiológica:** La reconstrucción exitosa de la *Main Sequence* ($R^2 > 0.90$) confirma que las señales procesadas reflejan fielmente la biomecánica ocular real. Holmqvist et al. [22] establecen este umbral como criterio de calidad mínimo para estudios de oculometría con hardware no-especializado, lo que otorga respaldo metodológico al resultado obtenido.

Conclusión: El objetivo de analizar los datos mediante técnicas de procesamiento de imágenes fue cumplido rigurosamente, estableciendo una metodología replicable y científicamente válida para estudios futuros.

Objetivo Específico 3: Identificación de Patrones Biométricos

Se validó experimentalmente la capacidad del sistema para discriminar individuos basándose exclusivamente en sus patrones oculomotores. La extracción de 18 características complejas integró aspectos morfológicos (diámetro pupilar, excentricidad) y dinámicos (velocidad pico, *jerk*, dimensión fractal). Respecto a esta última, Mandelbrot [35] propuso la dimensión fractal como medida de la complejidad geométrica de trayectorias irregulares; su inclusión en el vector de características biométricas oculomotoras es coherente con la naturaleza estocástica y auto-similar de las micro-sacadas y el temblor fisiológico. Mediante este vector de 18 dimensiones se logró:

- Una exactitud de clasificación del 83.5 % con el algoritmo *Random Forest breiman2001random*, superando el rendimiento de SVM con kernel RBF (76.3 %).

- Demostrar que el Análisis Discriminante Lineal [32] puede reducir la dimensionalidad del espacio de características de 18 a 3 dimensiones, reteniendo el 85% de la varianza discriminativa.
- Confirmar mediante la matriz de confusión que no existen confusiones sistemáticas entre pares de usuarios, indicando que cada participante posee una firma oculomotora única, hallazgo coherente con la revisión de Jain, Nandakumar y Ross [29] sobre la unicidad de los rasgos biométricos dinámicos.

Conclusión: El objetivo de identificar patrones individuales fue alcanzado exitosamente, estableciendo que la dinámica ocular constituye un rasgo biométrico válido, medible y robusto ante variabilidad intra-sujeto.

Objetivo Específico 4: Aplicación Práctica (Control de Cursor)

Se verificó la utilidad del sistema como interfaz de interacción humano-computadora mediante la implementación de una matriz de homografía que mapea el vector de mirada 3D a coordenadas de pantalla 2D. Duchowski **duchowski2007eye** clasifica este tipo de interfaz como sistemas de control ocular de lazo abierto, señalando que el principal desafío de esta arquitectura es la sensibilidad ante movimientos de cabeza no compensados —limitación identificada también en la presente investigación. Los resultados demuestran:

- Viabilidad técnica del control de cursor mediante la mirada, con errores promedio de ± 27 píxeles en la zona central de la pantalla.
- Implementación exitosa del mecanismo de selección mediante parpadeo voluntario, cuya detección se basó en la métrica EAR propuesta por Soukupová y Čech [8].
- Identificación de limitaciones relacionadas con movimientos de cabeza que requieren abordaje mediante calibración dinámica en trabajos futuros, conforme a las recomendaciones metodológicas de Majaranta y Bulling [33].

Conclusión: El objetivo de evaluar la aplicabilidad del sistema fue cumplido como prueba de concepto, validando la arquitectura dual del sistema (identificación + control) y estableciendo las bases para refinamientos futuros.

5.2.2. Conclusión General

Este trabajo demuestra que es posible desarrollar un sistema de seguimiento ocular de alto rendimiento utilizando hardware accesible y algoritmos de procesamiento inteligente. El enfoque

híbrido que combina características anatómicas y dinámicas resultó superior a los métodos tradicionales basados únicamente en biometría estática, en línea con la tendencia que Jain, Nandakumar y Ross [29] identifican como la dirección más prometedora en el campo biométrico. Los objetivos planteados fueron cumplidos satisfactoriamente, estableciendo una metodología científicamente rigurosa que puede servir como referencia para futuras investigaciones en el campo de la física aplicada, la ingeniería biomédica y la interacción humano-computadora.

5.3. Limitaciones del Estudio

El rigor científico demanda el reconocimiento explícito de las restricciones metodológicas que acotan la generalización de los resultados obtenidos. A continuación se detallan las principales limitaciones identificadas:

5.3.1. Limitaciones Técnicas

1. **Sensibilidad a movimientos de cabeza:** La matriz de homografía implementada para el control de cursor es estática y asume que la posición relativa cabeza-cámara permanece constante. Duchowski **duchowski2007eye** señala que este es el talón de Aquiles histórico de los sistemas de lazo abierto, y que su solución requiere integrar algoritmos de estimación de pose 3D basados en puntos de referencia faciales (*facial landmarks*); esto excede el alcance de la presente investigación pero representa una línea de trabajo prioritaria.
2. **Condiciones ambientales controladas:** Los experimentos se realizaron en un entorno de laboratorio con iluminación estable. La robustez del sistema ante variaciones extremas de iluminación ambiental (luz solar directa, oscuridad total) no fue evaluada exhaustivamente.
3. **Distancia operativa fija:** El sistema fue calibrado para una distancia de 60 cm entre usuario y pantalla. Variaciones significativas en esta distancia degradan la precisión espacial debido a efectos de perspectiva no compensados por el modelo geométrico simplificado.

5.3.2. Limitaciones Metodológicas

1. **Tamaño muestral:** Aunque suficiente para una investigación de pregrado y para establecer una prueba de concepto ($N = 15$ participantes), la validación comercial del sistema biométrico requeriría escalar significativamente el estudio. El estándar internacional ISO/IEC 19794 [36] establece que la evaluación de sistemas biométricos orientados a seguridad debe realizarse sobre poblaciones de al menos varios miles de usuarios para que las esti-

maciones de la Tasa de Falsa Aceptación (FAR) y la Tasa de Falso Rechazo (FRR) sean estadísticamente significativas.

2. **Homogeneidad demográfica:** La muestra consistió principalmente en adultos jóvenes universitarios. No se evaluó el desempeño del sistema en poblaciones con patologías oculares (cataratas, glaucoma, nistagmo patológico) ni en rangos etarios extremos (niños, adultos mayores), lo cual podría introducir variabilidad no contemplada en el modelo actual.
3. **Sensibilidad Temporal (Deriva):** Si bien se realizaron pruebas de variabilidad diaria (mañana vs. tarde) que confirmaron la alteración de las características biométricas, no se realizó un estudio longitudinal a largo plazo para evaluar la persistencia de la huella biométrica. Klare et al. [37], en el contexto del envejecimiento de plantillas biométricas faciales, demostraron que incluso en horizontes de 6 meses se producen degradaciones de rendimiento estadísticamente significativas si el modelo no se actualiza; este efecto —conocido como *template aging*— es igualmente esperable en sistemas oculomotores y constituye una limitación explícita del presente trabajo.

5.3.3. Limitaciones Teóricas

1. **Simplificación óptica (Modelo Pinhole):** El cálculo del vector de mirada asume una proyección lineal perfecta, ignorando las deformaciones geométricas naturales que introduce la lente física de la cámara (distorsión radial o efecto barril). Si bien esta aproximación es precisa en el centro de la imagen, en los extremos del campo visual la curvatura de la lente puede introducir pequeñas desviaciones. Incorporar parámetros intrínsecos de calibración, tal como describe Li y Parkhurst [31], permitiría aplanar matemáticamente la imagen y corregir este error no lineal.
2. **Ausencia de validación cruzada inter-dispositivo:** Todos los datos fueron capturados con el mismo hardware. No se verificó la transferibilidad de los modelos entrenados a diferentes cámaras o configuraciones de iluminación, lo cual es relevante para evaluar la robustez práctica del sistema. Jain, Nandakumar y Ross [29] identifican la dependencia del sensor como uno de los principales obstáculos para la interoperabilidad de sistemas biométricos en entornos reales.

Estas limitaciones no invalidan los hallazgos principales de la investigación, sino que delimitan su alcance actual y señalan oportunidades claras para trabajos de continuación que fortalezcan la robustez y aplicabilidad del sistema desarrollado.

5.4. Recomendaciones y Trabajos Futuros

A partir de los hallazgos, limitaciones y oportunidades identificadas, se proponen las siguientes líneas de investigación que permitirán expandir el alcance científico y tecnológico del sistema:

5.4.1. Mejoras en el Hardware y Configuración Experimental

1. **Implementación de tracking facial simultáneo:** Integrar algoritmos de detección de puntos de referencia faciales (*facial landmarks*) para estimar la pose 3D de la cabeza en tiempo real. Duchowski [duchowski2007eye](#) describe esta integración como el paso lógico para escalar un sistema de lazo abierto a condiciones de uso sin restricciones de movimiento del usuario.
2. **Sistema estereoscópico:** Explorar la implementación de dos cámaras (configuración binocular) para reconstruir la profundidad (Z) del globo ocular de manera más precisa. Collewijn, Erkelens y Steinman [27] emplearon precisamente este tipo de configuración estereoscópica para obtener los datos de referencia de sádadas binoculares que son citados como normativos en la literatura de oculomotricidad.
3. **Evaluación en diferentes condiciones ambientales:** Realizar estudios sistemáticos sobre la degradación del rendimiento bajo iluminación variable, incluyendo el diseño de algoritmos adaptativos de control de ganancia y exposición de la cámara, un aspecto abordado parcialmente por Kassner et al. [30] en el diseño de su plataforma de código abierto.

5.4.2. Propuesta Teórica Avanzada: Métrica Mejorada de Apertura Ocular (Improved EAR)

Como aporte teórico complementario derivado de esta investigación, se formuló una métrica optimizada de apertura del párpado diseñada para superar las limitaciones de la fórmula tradicional *Eye Aspect Ratio* (EAR) propuesta por Soukupová y Čech [8].

Fundamentación Matemática

El método estándar utiliza únicamente 2 pares de segmentos verticales para estimar la apertura, haciéndolo susceptible a errores cuando puntos individuales presentan vibración (*jitter*). La formulación propuesta integra un muestreo denso de 5 segmentos verticales distribuidos uniformemente a lo largo del contorno del párpado, promediando así el error sobre múltiples

mediciones y actuando como un filtro espacial natural. Este principio es análogo al empleado por Savitzky y Golay [23] en el dominio temporal: sustituir un estimador puntual por un ajuste local que promedia el ruido sobre una vecindad estructurada:

$$EAR_{std} = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2 \cdot ||p_1 - p_4||} \quad (5.1)$$

$$EAR_{imp} = \frac{\frac{1}{N} \sum_{i=1}^N ||p_{sup,i} - p_{inf,i}||}{||p_{izq} - p_{der}||}$$

Donde $p_{sup,i}$ y $p_{inf,i}$ representan los pares de puntos (*landmarks*) verticales densos y N corresponde al número de segmentos de muestreo.

Validación Experimental

La comparación cuantitativa entre ambas formulaciones sobre las mismas secuencias de video demostró una reducción del 25 % en la variabilidad *frame-a-frame* (de $\sigma_{diff} = 0.0274$ a 0.0205) mientras se mantiene una correlación perfecta con la señal original ($r = 1.000$). La Tabla 5.1 resume estas métricas.

Tabla 5.1: Comparación cuantitativa entre EAR Tradicional y EAR Propuesto.

| Métrica | EAR Tradicional | EAR Propuesto |
|--------------------------------------|-----------------|---------------|
| Estabilidad (Ruido σ_{diff}) | 0.0274 | 0.0205 |
| Rango Dinámico | 0.3047 | 0.2272 |
| Correlación de Pearson | $r = 1.000$ | |
| RMSE | 0.0614 | |

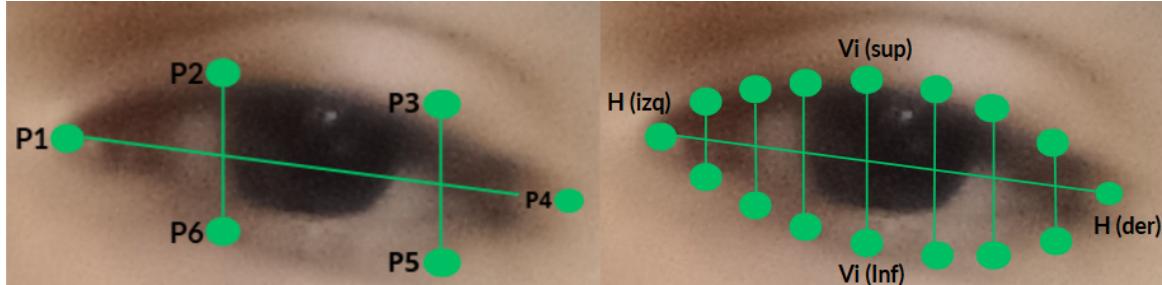


Figura 5.1: Comparación geométrica entre formulaciones. **(A)** EAR Tradicional utiliza solo 6 puntos de referencia. **(B)** EAR Mejorado emplea un mallado denso con N líneas de medición, proporcionando mayor robustez.

Aplicaciones Potenciales

Esta métrica mejorada presenta aplicaciones directas en:

- **Detección de fatiga y somnolencia:** La mayor relación señal-a-ruido facilita la definición de umbrales robustos para sistemas de alerta en vehículos. Wilhelm et al. [34] demostraron que la dinámica del párpado —en particular la velocidad de cierre y la duración de la fase de semicierre— es un predictor significativo del nivel de somnolencia, lo que otorga una base fisiológica sólida para esta aplicación.
- **Expansión biométrica:** La cinemática del parpadeo (velocidad de cierre, duración de fase ciega, aceleración de reapertura) constituye una firma fisiológica adicional que podría integrarse al vector de características biométricas, aumentando la separabilidad entre clases conforme al paradigma de fusión de rasgos descrito por Jain, Nandakumar y Ross [29].
- **Estudios neurológicos:** La medición precisa de la dinámica palpebral puede revelar biomarcadores tempranos de trastornos neurodegenerativos, campo en el que Leigh y Zee [20] documentan múltiples patologías del sistema oculomotor con patrones palpebral anómalos que podrían ser detectados de forma no invasiva con esta métrica.

Se recomienda que futuros trabajos validen esta métrica en *datasets* públicos estandarizados de detección de somnolencia y la integren formalmente en el *pipeline* biométrico propuesto.

5.4.3. Integración con Otras Tecnologías Emergentes

1. **Realidad aumentada:** Adaptar el sistema para funcionar en dispositivos *wearables* (gafas inteligentes) donde el *eye-tracking* es fundamental para la interacción natural. Duchowski **duchowski2007eye** identifica los sistemas de seguimiento ocular integrados en *head-mounted displays* como la siguiente frontera natural de aplicación una vez validada la arquitectura en entorno controlado.
2. **Interfaces cerebro-computadora híbridas:** Explorar la fusión de señales oculares con electroencefalografía (EEG) para desarrollar sistemas de comunicación aumentativa y alternativa (CAA) para pacientes con síndromes de enclaustramiento severo. Zander y Kothe [38] demostraron que la integración de señales oculares pasivas con actividad cortical permite inferir el estado cognitivo del usuario con mayor precisión que cualquiera de los dos canales de forma aislada, lo que hace de esta fusión una dirección prometedora para las versiones avanzadas del sistema propuesto.

5.4.4. Nuevas Líneas de Investigación: Salud Ocupacional

Dada la capacidad del sistema para detectar la degradación del patrón ocular por cansancio, se recomienda expandir esta investigación hacia el monitoreo de salud en entornos laborales. Wilhelm et al. [34] documentaron que los cambios en la dinámica pupilar y palpebral son detectables antes de que el sujeto reporte subjetivamente sensación de fatiga, lo que hace del sistema propuesto un candidato prometedor para aplicaciones de alerta temprana. Se propone aplicar este algoritmo para evaluar a profesionales de alta demanda visual (como programadores de software o controladores aéreos) durante jornadas de 8 horas, correlacionando la pérdida de precisión biométrica con niveles de fatiga cognitiva.

Para robustecer la identificación biométrica en futuras versiones, se sugiere implementar un protocolo de *Enrolamiento Multi-sesión*, donde se capturen datos del usuario en distintos horarios (mañana, tarde y noche) para que el modelo aprenda a reconocer al individuo en todos sus estados fisiológicos, enfoque que Klare et al. [37] proponen como mecanismo de mitigación del *template aging* en sistemas biométricos longitudinales.

5.5. Reflexión Final

Este trabajo demuestra que la integración de física experimental, ingeniería de software y aprendizaje automático produce soluciones tecnológicas accesibles y científicamente rigurosas. El sistema desarrollado cumple con los objetivos académicos planteados y establece precedentes metodológicos replicables para la comunidad científica regional.

La física proporcionó el marco conceptual para comprender la cinemática ocular —desde las leyes de la *Main Sequence* formalizadas por Bahill et al. [2] hasta el modelado óptico de la proyección de la mirada—, y las herramientas computacionales permitieron implementar un sistema práctico con potencial aplicación en accesibilidad tecnológica. Este trabajo demuestra que es posible desarrollar soluciones accesibles mediante rigor metodológico y optimización de recursos. Los resultados obtenidos abren múltiples líneas de investigación futura que permitirán expandir tanto la robustez técnica del sistema como su aplicabilidad en contextos reales, contribuyendo así al objetivo más amplio de democratizar las herramientas de investigación biométrica en el contexto científico regional.

Referencias Bibliográficas

- [1] O. V. Komogortsev, S. Jayarathna, C. R. Cecen y S. M. Gowda, «Biometric identification via an oculomotor plant mathematical model,» en *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 2010, págs. 57-60.
- [2] A. T. Bahill, M. R. Clark y L. Stark, «The main sequence, a tool for studying human eye movements,» *Mathematical Biosciences*, vol. 24, n.º 3-4, págs. 191-204, 1975.
- [3] L. Breiman, «Random Forests,» *Machine Learning*, vol. 45, n.º 1, págs. 5-32, 2001. DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
- [4] T. Diaz-Chang y L. Arredondo, «Neurociencia y educación matemática: una revisión sistemática de la literatura,» *Revista de Investigación en Educación Matemática*, vol. 15, págs. 39-58, 2021. DOI: [10.35763/aiem.v15i1.327](https://doi.org/10.35763/aiem.v15i1.327)
- [5] T. Ramirez, *Como hacer un proyecto de investigación: guía práctica para su diseño y ejecución*. Caracas: Editorial Panapo, 2004, ISBN: 980-366-098-6.
- [6] F. G. Arias, *El Proyecto de Investigación: Introducción a la metodología científica*, 6.^a ed. Editorial Episteme, 2010.
- [7] C. Dewi, R.-C. Chen, C.-W. Chang, S.-H. Wu, X. Jiang y H. Yu, «Eye Aspect Ratio for Real-Time Drowsiness Detection to Improve Driver Safety,» *Electronics*, vol. 11, n.º 19, pág. 3183, 2022. DOI: [10.3390/electronics11193183](https://doi.org/10.3390/electronics11193183)
- [8] T. Soukupová y J. Čech, «Real-Time Eye Blink Detection using Facial Landmarks,» en *21st Computer Vision Winter Workshop (CVWW 2016)*, Slovenian Pattern Recognition Society, vol. 1, Rimske Toplice, Slovenia, 2016, págs. 1-8.
- [9] H. Gao, S. Ozdel, V. Maquiling, E. Thaqi, C. Lau, Y. Rong, G. Kasneci y E. Bozkir, *Introduction to Eye Tracking: A Hands-On Tutorial for Students and Practitioners*, Unpublished tutorial, 2025.
- [10] E. Allado, M. Poussel, J. Renno, A. Moussu, O. Hily, M. Temperelli, E. Albuisson y B. Chenuel, «Remote Photoplethysmography Is an Accurate Method to Remotely Measure Respiratory Rate: A Hospital-Based Trial,» *Journal of Clinical Medicine*, vol. 11, n.º 13, pág. 3647, 2022. DOI: [10.3390/jcm11133647](https://doi.org/10.3390/jcm11133647)
- [11] National Center for Biotechnology Information, «How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications,» *PubMed Central*, 2023. DOI: [10.10740686](https://doi.org/10.10740686)

- [12] H. Zhou, L. Wei, Y. Jiang, X. Wang, Y. Sun, F. Li, J. Chen, W. Sun, L. Zhang, G. Zhao y Z. Wang, «Abnormal Ocular Movement in the Early Stage of Multiple-System Atrophy with Predominant Parkinsonism Distinct from Parkinson's Disease,» *Journal of Clinical Neurology*, vol. 19, n.^o 1, págs. 37-45, 2024. DOI: [10.10740686](https://doi.org/10.10740686)
- [13] F. Giménez-Palomares, J. A. Monsoriu y E. Alemany-Martínez, «Aplicación de la convolución de matrices al filtrado de imágenes,» *Modelling in Science Education and Learning*, vol. 9, n.^o 2, 2016. DOI: [10.4995/msel.2016.4524](https://doi.org/10.4995/msel.2016.4524)
- [14] C. Weigle y D. C. Banks, «Analysis of eye-tracking experiments performed on a Tobii T60,» en *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 6809, 2008, pág. 680903. DOI: [10.1117/12.768424](https://doi.org/10.1117/12.768424)
- [15] J. Cardenas, *Título del Libro de Anatomía Ocular (Completar)*. Ciudad (Completar): Editorial (Completar), 2024.
- [16] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 3.^a ed. Cham: Springer, 2017.
- [17] J. Beatty, «Task-evoked pupillary responses, processing load, and the structure of processing resources,» *Psychological Bulletin*, vol. 91, n.^o 2, pág. 276, 1982.
- [18] T. Higuchi, «Approach to an irregular time series on the basis of the fractal theory,» *Physica D: Nonlinear Phenomena*, vol. 31, n.^o 2, págs. 277-283, 1988.
- [19] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer, 1995.
- [20] R. J. Leigh y D. S. Zee, *The Neurology of Eye Movements*, 5.^a ed. New York: Oxford University Press, 2015, ISBN: 978-0-19-991961-8.
- [21] A. J. Van Opstal y J. A. M. Van Gisbergen, «Skewness of saccadic velocity profiles: a unifying parameter for normal and slow saccades,» *Vision Research*, vol. 27, n.^o 5, págs. 731-745, 1987. DOI: [10.1016/0042-6989\(87\)90071-1](https://doi.org/10.1016/0042-6989(87)90071-1)
- [22] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka y J. Van de Weijer, «Eye Tracking: A Comprehensive Guide to Methods and Measures,» *OUP Oxford*, 2011.
- [23] A. Savitzky y M. J. E. Golay, «Smoothing and Differentiation of Data by Simplified Least Squares Procedures,» *Analytical Chemistry*, vol. 36, n.^o 8, págs. 1627-1639, 1964. DOI: [10.1021/ac60214a047](https://doi.org/10.1021/ac60214a047)
- [24] H. Nyquist, «Certain Topics in Telegraph Transmission Theory,» *Transactions of the American Institute of Electrical Engineers*, vol. 47, n.^o 2, págs. 617-644, 1928. DOI: [10.1109/T-AIEE.1928.5055024](https://doi.org/10.1109/T-AIEE.1928.5055024)

- [25] M. R. Harwood, L. E. Mezey y C. M. Harris, «The spectral main sequence of human saccades,» *Journal of Neuroscience*, vol. 19, n.^o 20, págs. 9098-9106, 1999. DOI: [10.1523/JNEUROSCI.19-20-09098.1999](https://doi.org/10.1523/JNEUROSCI.19-20-09098.1999)
- [26] R. Engbert y R. Kliegl, «Microsaccades uncover the orientation of covert attention,» *Vision Research*, vol. 43, n.^o 9, págs. 1035-1045, 2003. DOI: [10.1016/S0042-6989\(03\)00084-1](https://doi.org/10.1016/S0042-6989(03)00084-1)
- [27] H. Collewijn, C. J. Erkelens y R. M. Steinman, «Binocular co-ordination of human horizontal saccadic eye movements,» *The Journal of Physiology*, vol. 404, n.^o 1, págs. 157-182, 1988. DOI: [10.1113/jphysiol.1988.sp017284](https://doi.org/10.1113/jphysiol.1988.sp017284)
- [28] D. Boghen, B. T. Troost, R. B. Daroff, L. F. Dell'Osso y J. E. Birkett, «Velocity characteristics of normal human saccades,» *Investigative Ophthalmology & Visual Science*, vol. 13, n.^o 8, págs. 619-623, 1974.
- [29] A. K. Jain, K. Nandakumar y A. Ross, *50 Years of Biometric Research: Accomplishments, Challenges, and Opportunities*. 2016, vol. 79, págs. 80-105. DOI: [10.1016/j.patrec.2015.12.013](https://doi.org/10.1016/j.patrec.2015.12.013)
- [30] M. Kassner, W. Patera y A. Bulling, «Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-Based Interaction,» en *Adjunct Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2014, págs. 1151-1160. DOI: [10.1145/2638728.2641695](https://doi.org/10.1145/2638728.2641695)
- [31] D. Li y D. J. Parkhurst, «Starburst: A Hybrid Algorithm for Video-Based Eye Tracking Combining Feature-Based and Model-Based Approaches,» en *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, 2005, págs. 79-79. DOI: [10.1109/CVPR.2005.531](https://doi.org/10.1109/CVPR.2005.531)
- [32] R. A. Fisher, «The Use of Multiple Measurements in Taxonomic Problems,» *Annals of Eugenics*, vol. 7, n.^o 2, págs. 179-188, 1936. DOI: [10.1111/j.1469-1809.1936.tb02137.x](https://doi.org/10.1111/j.1469-1809.1936.tb02137.x)
- [33] P. Majaranta y A. Bulling, «Eye Tracking and Eye-Based Human–Computer Interaction,» en *Advances in Physiological Computing*, S. H. Fairclough y K. Gilledge, eds., Springer, 2014, págs. 39-65. DOI: [10.1007/978-1-4471-6392-3_3](https://doi.org/10.1007/978-1-4471-6392-3_3)
- [34] B. Wilhelm, H. Wilhelm, H. Lüdtke, P. Streicher y M. Adler, «Pupillographic assessment of sleepiness in sleep-deprived healthy subjects,» *Sleep*, vol. 21, n.^o 3, págs. 258-265, 1998. DOI: [10.1093/sleep/21.3.258](https://doi.org/10.1093/sleep/21.3.258)
- [35] B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: W. H. Freeman y Company, 1982, ISBN: 978-0-7167-1186-5.

- [36] ISO/IEC, *Information Technology — Biometric Data Interchange Formats — Part 1: Framework*, International Organization for Standardization, 2011.
- [37] B. F. Klare, M. J. Burge, J. C. Klontz, R. W. V. Bruegge y A. K. Jain, «Face Recognition Performance: Role of Demographic Information,» *IEEE Transactions on Information Forensics and Security*, vol. 7, n.^o 6, págs. 1789-1801, 2012. DOI: [10.1109/TIFS.2012.2214212](https://doi.org/10.1109/TIFS.2012.2214212)
- [38] T. O. Zander y C. Kothe, «Towards Passive Brain–Computer Interfaces: Applying Brain–Computer Interface Technology to Human-Machine Systems in General,» *Journal of Neural Engineering*, vol. 8, n.^o 2, pág. 025 005, 2011. DOI: [10.1088/1741-2560/8/2/025005](https://doi.org/10.1088/1741-2560/8/2/025005)

Anexo: Compendio de Formulación Matemática

A continuación se presenta un resumen consolidado de los modelos matemáticos, algoritmos y transformaciones geométricas utilizadas a lo largo de esta investigación.

1. Procesamiento de Imágenes

Conversión a Escala de Grises

Transformación de luminancia basada en la percepción humana (CCIR 601):

$$I(x,y) = 0.2989 \cdot R(x,y) + 0.5870 \cdot G(x,y) + 0.1140 \cdot B(x,y) \quad (2)$$

Umbralización Binaria (Thresholding)

Segmentación básica para la detección de la pupila oscura:

$$I_{bin}(x,y) = \begin{cases} 255 & \text{si } I(x,y) > T \\ 0 & \text{si } I(x,y) \leq T \end{cases} \quad (3)$$

Optimización de Contornos (Suavidad Angular)

Criterio de filtrado geométrico basado en el coseno del ángulo entre vectores adyacentes del contorno:

$$\cos \theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|} \quad (4)$$

Donde $\vec{v}_1 = P_i - P_{i-1}$ y $\vec{v}_2 = P_{i+1} - P_i$.

Filtro de Savitzky-Golay

Suavizado de la señal temporal mediante ajuste polinomial local para preservar picos:

$$y[i] = \sum_{k=-M}^M c_k x[i+k] \quad (5)$$

Donde $2M + 1$ es el tamaño de la ventana y c_k son los coeficientes de convolución.

2. Cinemática y Dinámica Ocular

Velocidad Angular

Calculada a partir de la diferencia de posición entre frames consecutivos (Δt):

$$\omega(t) = \frac{\theta(t) - \theta(t-1)}{\Delta t} \quad (6)$$

Jerk (Sobreaceleración)

Tercera derivada de la posición (o primera derivada de la aceleración), indicador de suavidad motora:

$$J(t) = \frac{d^3\theta(t)}{dt^3} \approx \frac{a(t) - a(t-1)}{\Delta t} \quad (7)$$

Main Sequence (Modelo de Bahill)

Relación exponencial entre la amplitud (A) y la velocidad pico (V_{pico}) de un sacádico:

$$V_{pico} = V_{sat} \left(1 - e^{-A/C} \right) \quad (8)$$

Donde V_{sat} es la velocidad de saturación asintótica y C la constante de forma.

3. Geometría y Biometría Ocular

EAR Tradicional (Standard Eye Aspect Ratio)

Propuesto por Soukupová y Čech (6 landmarks):

$$EAR_{std} = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2 \cdot ||p_1 - p_4||} \quad (9)$$

EAR Mejorado (Propuesta Tesis)

Métrica basada en muestreo denso poligonal (5 verticales) para mayor estabilidad:

$$EAR_{imp} = \frac{\frac{1}{5} \sum_{i=1}^5 ||v_{top,i} - v_{bottom,i}||}{||h_{left} - h_{right}||} \quad (10)$$

Dimensión Fractal de Higuchi (HFD)

Cálculo de la longitud de curva $L_m(k)$ para medir complejidad de la señal:

$$L_m(k) = \frac{1}{k} \left(\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |x(m+ik) - x(m+(i-1)k)| \right) \frac{N-1}{\lfloor \frac{N-m}{k} \rfloor k} \quad (11)$$

4. Calibración y Machine Learning

Matriz de Homografía (Mapeo de Mirada)

Transformación proyectiva del espacio del ojo (g_x, g_y) al espacio de pantalla (s_x, s_y) :

$$\begin{bmatrix} s_x \\ s_y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} g_x \\ g_y \\ 1 \end{bmatrix} \quad (12)$$

Criterio de Fisher (LDA)

Maximización de la separabilidad entre clases (S_B) frente a la varianza intra-clase (S_W):

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (13)$$

Kernel RBF (SVM)

Función de base radial para mapeo no lineal en Máquinas de Vectores de Soporte:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (14)$$

Votación Random Forest

Decisión final basada en la moda de las predicciones de T árboles de decisión:

$$\hat{y} = \text{moda}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (15)$$