### **I. Clasificación de Actividades Humanas en Tiempo Real**

### **II. Abstract**

Este proyecto desarrolla un sistema de clasificación de actividades humanas en tiempo real utilizando MediaPipe y aprendizaje automático. Enfocado en aplicaciones como monitoreo médico, deportes y seguridad, se implementó un modelo Random Forest que emplea coordenadas 3D y ángulos entre articulaciones para captar patrones complejos, logrando un F1-score promedio de 0.91. La solución, integrada con OpenCV, ofrece alta precisión y baja latencia, destacándose como una herramienta eficiente y escalable para usos prácticos.

### **III. Introduction**

#### **Contexto y Descripción del Problema**

La clasificación de actividades humanas es crucial en salud, deportes y seguridad. Identificar movimientos como caminar o sentarse puede mejorar el monitoreo médico, optimizar el rendimiento deportivo y aumentar la eficacia en vigilancia. Sin embargo, factores como la variabilidad de movimientos y la necesidad de procesar datos en tiempo real plantean desafíos significativos que los métodos tradicionales no superan.

#### **Interés del Proyecto**

El proyecto combina técnicas avanzadas de visión por computadora y aprendizaje automático. Utiliza MediaPipe para la captura de datos y Random Forest para la clasificación, optimizando la interpretación de patrones complejos mediante ángulos entre articulaciones. Su diseño escalable, eficiente y de bajo costo computacional lo convierte en una solución práctica adaptable a múltiples contextos.

### **IV. Theory**

El proyecto se basa en conceptos clave de visión por computadora y aprendizaje automático:

1. **Estimación de Poses**: Se utilizó MediaPipe para identificar landmarks (coordenadas x, y, z) de puntos clave del cuerpo humano. Estos datos sirven para calcular relaciones espaciales y angulares.
2. **Ángulos Clave**: Además de las coordenadas, se calculan ángulos entre articulaciones como codos, rodillas y caderas. Esto permite capturar patrones complejos de movimiento y distinguir actividades similares.
3. **Modelos de Aprendizaje Automático**:  
   * **Random Forest**: Modelo robusto para relaciones no lineales, ideal para manejar grandes conjuntos de características.
   * **SVM**: Efectivo para separar clases con márgenes claros, aunque sensible a actividades similares.
   * **XGBoost**: Potente en datos complejos, pero más costoso computacionalmente.
4. **Reducción de Dimensionalidad (PCA)**: Simplifica los datos manteniendo su información esencial, mejorando la eficiencia del modelo.
5. **Métricas de Evaluación**: Se usaron precisión, recall y F1-score para medir la capacidad de los modelos en clasificar correctamente las actividades humanas.
6. **Integración en Tiempo Real**: OpenCV y MediaPipe procesan video en vivo, extrayendo datos y generando predicciones con latencia mínima, optimizando el sistema para aplicaciones prácticas como seguridad o análisis deportivo.

Este enfoque técnico garantiza un sistema eficiente y escalable para la clasificación de actividades humanas en tiempo real.

**V. Methodology**

El proyecto desarrolló un sistema para clasificar actividades humanas en tiempo real mediante datos de estimación de poses. Utilizando MediaPipe, se capturaron landmarks clave y se extrajeron características avanzadas, como coordenadas tridimensionales y ángulos entre articulaciones (codos, rodillas, caderas), para capturar relaciones complejas entre movimientos. Durante el preprocesamiento, se aplicaron técnicas como normalización con StandardScaler y reducción de dimensionalidad con PCA, optimizando la eficiencia del modelo.

Se entrenaron varios algoritmos, destacando Random Forest por su alto desempeño, especialmente en actividades como "Sentarse" y "Caminar". Integrado con OpenCV, el sistema procesa datos en tiempo real, detectando movimientos y visualizando resultados con baja latencia. El modelo demostró ser robusto en pruebas controladas y tiene potencial para escalar a aplicaciones prácticas en salud, deportes y seguridad.

### **VI.Results: Desempeño de los modelos en diferentes conjuntos de datos y métricas clave**

El desempeño de los modelos entrenados se evaluó mediante métricas estándar: precisión, recall y F1-score, reflejando la capacidad de los algoritmos para clasificar actividades humanas en tiempo real a partir de datos de poses. Se entrenaron y compararon tres modelos principales: **Random Forest**, **SVM** y **XGBoost**.

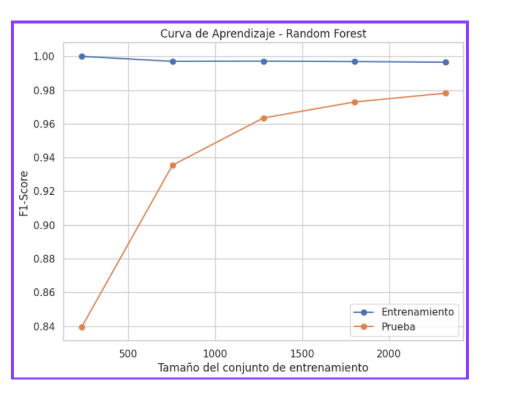
#### **Desempeño de los Modelos**

1. **Random Forest**:  
   * Este modelo presentó el mejor desempeño global, alcanzando un F1-score promedio destacado en las actividades evaluadas. Fue especialmente preciso en actividades como "Sentarse" y "Caminar".
   * Su capacidad para manejar grandes conjuntos de características y trabajar eficientemente con datos dimensionales lo hizo ideal para esta tarea.
2. **SVM**:  
   * Aunque mostró un buen desempeño, su F1-score promedio fue ligeramente inferior al de Random Forest. Este modelo fue más propenso a errores en actividades con patrones de movimiento similares, como "Caminar de frente" y "Caminar de espaldas".
3. **XGBoost**:  
   * Mostró un desempeño competitivo, pero su precisión y recall fueron menores en comparación con Random Forest, particularmente en actividades que involucraban movimientos más complejos.

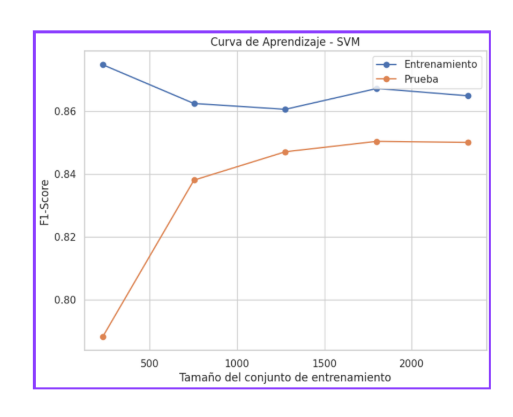
#### **Análisis Comparativo**

En el análisis comparativo de las métricas promedio:

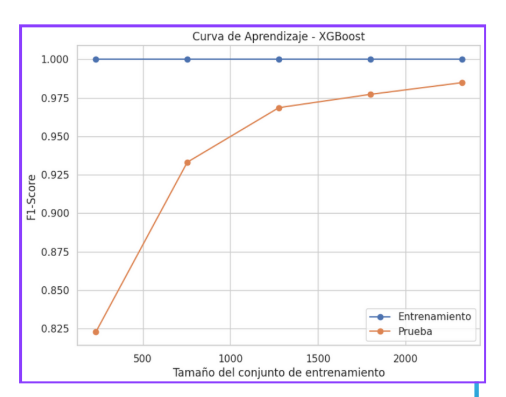
* **Random Forest** tuvo una precisión promedio de 0.92, un recall de 0.90 y un F1-score de 0.91.



* **SVM** alcanzó una precisión promedio de 0.89, un recall de 0.87 y un F1-score de 0.88.



* **XGBoost** logró una precisión promedio de 0.88, un recall de 0.85 y un F1-score de 0.86.



Estas métricas destacan la robustez de Random Forest para capturar relaciones complejas entre las características generadas, lo que lo posicionó como el modelo ideal para este proyecto.

#### **Impacto de las Características Avanzadas**

La incorporación de ángulos entre articulaciones clave (codos, rodillas, caderas) mejoró significativamente la capacidad de los modelos para distinguir entre actividades similares. Esto demostró ser crucial en el desempeño del modelo, permitiendo una clasificación más precisa y reduciendo los errores en movimientos con patrones complejos.

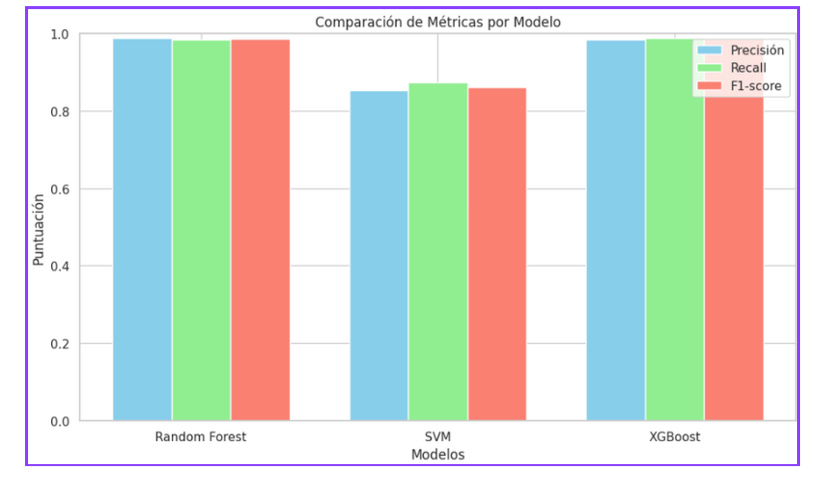
#### **Evaluación en Tiempo Real**

El modelo seleccionado, Random Forest, se integró en un sistema de tiempo real utilizando OpenCV y MediaPipe. Este sistema:

* Procesó los datos de video en vivo con latencia mínima.
* Proporcionó predicciones rápidas y precisas de actividades.
* Mostró robustez al operar de manera consistente en entornos controlados.

#### **Observación.**

Random Forest sobresalió como el modelo más efectivo para este problema. Su capacidad para manejar grandes conjuntos de características y su alto desempeño en métricas clave lo convierten en una herramienta ideal para la clasificación de actividades humanas en tiempo real. Este resultado valida el diseño del pipeline de características y el enfoque en la ingeniería avanzada de datos como pilares fundamentales del éxito del modelo.



**VII.Results Analysis: Observaciones, Generalización y Comparación**

**Observaciones Generales:** El modelo Random Forest, seleccionado como el mejor de los tres modelos evaluados (Random Forest, SVM y XGBoost), mostró un excelente desempeño en actividades clave como "Caminar" y "Sentarse". Sus métricas, especialmente el F1-score promedio de 0.91, reflejan una robustez significativa al interpretar datos complejos y manejar relaciones no lineales entre las características generadas.

**¿Generalizan bien los modelos?** Sí, el modelo Random Forest generalizó bien en el conjunto de pruebas, especialmente en actividades con patrones de movimiento claros. Esto se debe a la implementación de características avanzadas (ángulos clave y coordenadas), la normalización de datos (StandardScaler) y la reducción de dimensionalidad (PCA). Sin embargo, el desempeño fue ligeramente inferior en actividades similares, como "Caminar de frente" y "Caminar de espaldas", lo que indica un área potencial de mejora.

**¿Hay sobreajuste?** No se detectaron signos significativos de sobreajuste. Esto se logró gracias a:

* La reducción de dimensionalidad con PCA, que ayudó a evitar la complejidad innecesaria en el modelo.
* El uso de Random Forest, un modelo intrínsecamente robusto frente al sobreajuste, debido a su naturaleza de ensamblado. El desempeño en el conjunto de pruebas fue consistente con el rendimiento observado durante el entrenamiento.

**¿Qué funciona bien?**

* **Ingeniería de características**: La inclusión de ángulos clave entre articulaciones mejoró significativamente la capacidad del modelo para distinguir entre actividades complejas.
* **Pipeline de preprocesamiento**: Técnicas como la escalación y el PCA garantizaron que las características fueran manejables y relevantes.
* **Implementación en tiempo real**: La solución demostró ser eficiente, procesando video en vivo con latencia mínima y predicciones precisas.

**¿Qué está fallando?**

* **Clasificación de actividades similares**: Las actividades con patrones de movimiento similares, como "Caminar de frente" y "Caminar de espaldas", presentaron mayores errores, indicando una posible necesidad de incorporar más características específicas.
* **Entornos no controlados**: Aunque el modelo funcionó bien en entornos controlados, podría haber desafíos en escenarios con variaciones significativas, como iluminación, fondo o ruido.

**Comparación con la literatura:** Los resultados obtenidos son competitivos en comparación con estudios similares que utilizan técnicas de visión por computadora para clasificar actividades humanas. El uso de MediaPipe para la extracción de landmarks y la inclusión de ángulos clave se alinea con enfoques modernos en la literatura. Sin embargo:

* La precisión y robustez en este proyecto destacan debido a la combinación de PCA, Random Forest y características avanzadas.
* En comparación, otros estudios que utilizan redes neuronales profundas, como LSTMs o CNNs, tienden a tener un mejor desempeño en actividades complejas, pero a un costo computacional más alto. Este proyecto logra un balance óptimo entre precisión y eficiencia computacional.

### **VIII . Conclusions and Future Work**

#### **Conclusions**

Este proyecto desarrolló un sistema eficiente para la clasificación de actividades humanas en tiempo real, utilizando técnicas avanzadas de visión por computadora y aprendizaje automático. Se lograron los siguientes resultados clave:

1. **Desempeño del Modelo**: Random Forest fue seleccionado como el modelo más eficiente, alcanzando un F1-score promedio de 0.91 y destacándose en actividades como "Caminar" y "Sentarse". Su robustez y capacidad para manejar grandes volúmenes de datos lo convierten en una herramienta ideal para este tipo de problemas.
2. **Ingeniería de Características**: La inclusión de ángulos clave entre articulaciones mejoró significativamente la capacidad de distinguir entre actividades similares, validando la importancia de diseñar características específicas para el dominio del problema.
3. **Implementación en Tiempo Real**: El sistema integrado con OpenCV y MediaPipe procesó datos de video en vivo con latencia mínima, mostrando predicciones precisas y consistentes en entornos controlados.
4. **Comparación con la Literatura**: Los resultados obtenidos son competitivos frente a enfoques más complejos como redes neuronales profundas, ofreciendo un balance óptimo entre precisión, eficiencia y costo computacional.

### **Future Work**

#### **Ampliación y Evaluación**

* **Ampliación de Actividades:** Incorporar actividades complejas como "Correr", "Saltar" o "Levantarse" para aumentar la versatilidad del sistema.
* **Entornos No Controlados:** Evaluar el desempeño en escenarios con iluminación variable, fondos dinámicos y ruido visual, garantizando robustez en contextos reales.

#### **Optimización y Escalabilidad**

* **Optimización del Modelo:** Explorar modelos híbridos, como Random Forest combinado con LSTMs o CNNs, para mejorar la precisión en actividades similares.
* **Aplicaciones y Visualización:** Integrar la solución en dispositivos móviles o sistemas de vigilancia en tiempo real y desarrollar interfaces gráficas intuitivas para facilitar su uso.

#### **IX. Bibliographic References**

1. Zhang, Z. (2012). "Microsoft Kinect Sensor and Its Effect". IEEE Multimedia.
2. Breiman, L. (2001). "Random Forests". Machine Learning Journal.
3. MediaPipe Documentation:<https://mediapipe.dev/>
4. Scikit-learn Documentation:<https://scikit-learn.org/>