

Estrategia para la Obtención de Nuevos Datos

1. Incremento del Dataset

Para mejorar la capacidad del modelo de generalizar y adaptarse a diferentes escenarios, se implementaron las siguientes estrategias:

- **Ampliación del Grupo de Participantes:** Se agregaron 5 nuevos individuos al proceso de recolección, seleccionados para incluir diversidad en términos de edad, género, altura y compleción física.
- **Recolección en Diferentes Condiciones:** Los datos fueron capturados bajo distintas condiciones de iluminación y fondos para reflejar mejor los escenarios reales.
- **Variación en las Actividades:** Se introdujeron variaciones en las actividades, como caminar a diferentes velocidades o realizar giros con diferentes ángulos, aumentando la riqueza y complejidad del dataset.

2. Generación de Datos Sintéticos

Para ampliar aún más el dataset, se aplicaron técnicas de augmentación:

- **Ruido Aleatorio:** Simulación de errores menores en las coordenadas para representar variaciones naturales.
- **Transformaciones Geométricas:** Alteraciones en los ángulos calculados para simular diferentes perspectivas.
- **Interpolación Temporal:** Generación de registros adicionales interpolando entre poses consecutivas.

Resultado Final: El dataset pasó de 1000 a **2500 registros**, proporcionando una base más diversa y representativa para el entrenamiento del modelo.

Preparación de Datos

Los datos recolectados y generados fueron preprocesados siguiendo estos pasos:

1. **Normalización de Coordenadas:** Las coordenadas x, y, z de los landmarks se normalizaron entre 0 y 1 para eliminar dependencias relacionadas con el tamaño físico de las personas o su distancia de la cámara.
2. **Cálculo de Ángulos Clave:** Se calcularon métricas adicionales utilizando trigonometría, como:
 - Ángulos de rodilla (izquierda y derecha).
 - Inclinación del tronco.
 - Rotación relativa de los hombros.
3. **Manejo de Datos Faltantes:**

- Se utilizaron interpolaciones lineales para completar registros con datos faltantes menores.
 - Los registros con múltiples landmarks ausentes fueron descartados.
4. **División del Dataset:**
- **80%** para entrenamiento.
 - **20%** para pruebas.
 - Se utilizó una división estratificada para asegurar que cada actividad estuviera representada de manera equitativa en ambos conjuntos.
-

Entrenamiento de los Modelos

Modelos Utilizados

1. **Random Forest:** Con su capacidad para manejar relaciones no lineales y alta dimensionalidad, este modelo fue una opción robusta para tareas iniciales.
2. **XGBoost:** Este modelo optimizado para boosting ofreció un rendimiento sobresaliente en la clasificación de actividades dinámicas.

Optimización de Hiperparámetros

Se utilizaron dos enfoques principales:

- **Grid Search:** Exploración exhaustiva de combinaciones de hiperparámetros en Random Forest.
- **Optuna:** Algoritmo de optimización bayesiana para XGBoost, permitiendo explorar espacios más grandes de manera eficiente.

Hiperparámetros Ajustados:

- **Random Forest:**
 - `n_estimators`: 200
 - `max_depth`: 20
 - `min_samples_split`: 5
 - `min_samples_leaf`: 2
 - **XGBoost:**
 - `max_depth`: 7
 - `learning_rate`: 0.1
 - `n_estimators`: 300
 - `subsample`: 0.9
 - `colsample_bytree`: 0.8
-

Resultados Obtenidos

1. Métricas de Evaluación

Ambos modelos se evaluaron con las siguientes métricas:

- **Precisión Global:**
 - Random Forest: **95.2%**
 - XGBoost: **97.8%**
- **Recall por Clase:**
 - Las actividades estáticas como "standing" alcanzaron un recall del **98%**.
 - Las dinámicas, como "turning", lograron un recall del **95%**.
- **F1-Score Promedio:**
 - Random Forest: **94.8%**
 - XGBoost: **96.7%**

2. Visualizaciones

- **Matriz de Confusión:** Una matriz de calor resaltó la alta precisión de XGBoost, con pocas confusiones entre actividades como "walking towards" y "walking away".
 - **Importancia de Características:**
 - Los landmarks de caderas y rodillas resultaron ser los más influyentes en actividades dinámicas.
 - Los ángulos del tronco y hombros dominaron en actividades estáticas.
-

Plan de Despliegue

Uso de PyInstaller

Para distribuir el sistema, se decidió empaquetarlo en un archivo ejecutable independiente utilizando **PyInstaller**.

Ventajas de PyInstaller

1. **Portabilidad:** Permite que el sistema sea ejecutado en cualquier máquina sin necesidad de instalar Python ni sus dependencias.
2. **Compatibilidad Multiplataforma:** Soporte para Windows, macOS y Linux.
3. **Integración de Recursos:** Facilita la inclusión de archivos auxiliares como modelos entrenados, configuraciones y recursos gráficos.

Proceso de Empaquetado

1. **Preparación del Entorno:**

Instalación de PyInstaller:

```
pip install pyinstaller
```

- - Verificación de dependencias en el entorno virtual.
2. **Comando Principal:**

Para generar un archivo ejecutable:

```
pyinstaller --onefile --windowed main.py
```

○

Incluir archivos auxiliares (por ejemplo, modelos entrenados):

```
pyinstaller --onefile --add-data "pose_model.pkl:." main.py
```

○

3. **Pruebas del Ejecutable:**
- Validar el archivo generado en diferentes máquinas para asegurar que funcione correctamente.
-

Impacto del Proyecto

Impactos Positivos

1. **Eficiencia Operativa:** Automatización de tareas como el reconocimiento de actividades en tiempo real.
2. **Seguridad y Prevención:** Monitoreo proactivo de comportamientos potencialmente peligrosos.
3. **Escalabilidad:** El sistema es adaptable para incluir nuevas actividades o contextos.

Aspectos Éticos

1. **Privacidad:** Anonimizar los datos recolectados para proteger la identidad de los usuarios.
 2. **Equidad:** Garantizar que el modelo no discrimine entre diferentes grupos demográficos.
 3. **Transparencia:** Asegurar que los resultados del modelo sean interpretables y explicables.
-

Conclusión

El sistema alcanzó un desempeño sobresaliente con métricas robustas, demostrando su capacidad para clasificar actividades humanas de manera efectiva. El uso de PyInstaller asegura que sea portátil y accesible para una variedad de usuarios. A futuro, el enfoque

estará en expandir su aplicabilidad mientras se mantienen altos estándares éticos y técnicos.