

Reporte del Proyecto Final: Sistema de Anotación de Video

Abstract

Este proyecto tiene como objetivo desarrollar una herramienta capaz de analizar y clasificar actividades humanas específicas a través de la captura y procesamiento de video en tiempo real. Utilizando técnicas de inteligencia artificial y modelos de machine learning, se implementó un sistema que clasifica acciones como caminar, girar, sentarse y levantarse, basándose en el seguimiento de articulaciones clave y medidas posturales.

Introducción

La Universidad ICESI, a través de la Facultad de Ingeniería, Diseño y Ciencias Aplicadas, busca abordar problemas reales mediante el uso de modelos de analítica y conjuntos de datos de diversos formatos. Este proyecto se enfoca en la creación de un sistema de anotación de video que permita realizar un seguimiento de movimientos articulares y posturales en actividades cotidianas. La capacidad de clasificar estas acciones en tiempo real tiene aplicaciones en áreas como la rehabilitación física, el deporte y la ergonomía, donde el monitoreo preciso de la actividad es crucial.

Teoría

Para la comprensión de este desarrollo, es importante entender conceptos como:

- **Machine Learning (ML):** Un subcampo de la inteligencia artificial que se enfoca en la construcción de algoritmos que permiten a las computadoras aprender de los datos.
- **Support Vector Machine (SVM):** Un modelo supervisado utilizado para clasificación que busca encontrar el mejor margen entre diferentes clases en un espacio multidimensional.
- **Seguimiento de Articulaciones:** Métodos como MediaPipe permiten la detección de posiciones de articulaciones en tiempo real, lo que es esencial para el análisis de movimientos.

Metodología

El enfoque del proyecto se basa en la metodología CRISP-DM, que incluye las siguientes etapas:

1. **Recolección de Datos:** Videos fueron capturados con múltiples personas realizando diversas actividades. Se utilizó MediaPipe para el seguimiento de articulaciones.
2. **Preparación de los Datos:** Se realizaron procesos de limpieza, normalización y generación de características relevantes, como la velocidad de las articulaciones y los ángulos entre ellas.
3. **Entrenamiento del Modelo:** Se eligió el modelo SVM para la clasificación de actividades. Se dividieron los datos en conjuntos de entrenamiento y prueba y se realizó un ajuste de hiperparámetros.

4. **Evaluación del Modelo:** Se calcularon métricas como precisión, recall y F1-score para evaluar el rendimiento del modelo.

Preprocesamiento de Datos

El preprocesamiento es crucial para mejorar la calidad de los datos y la efectividad del modelo de clasificación. Se realizaron las siguientes etapas:

Normalización

Se estandarizaron las coordenadas de las articulaciones para evitar la dependencia de la altura de los sujetos o la distancia de la cámara. Esto asegura que las diferencias en la posición de la cámara no influyan en el desempeño del modelo.

Filtrado

Se aplicó un filtro suave a las posiciones de las articulaciones para eliminar el ruido generado durante el seguimiento. Esto ayuda a obtener trayectorias más limpias y precisas, lo que mejora la calidad de las características extraídas.

Generación de Características

Se extrajeron varias características útiles para el clasificador, incluyendo:

- **Velocidad de las articulaciones:** Se calcularon las velocidades de movimiento de las articulaciones a partir de las posiciones en diferentes frames.
- **Ángulos relativos entre articulaciones:** Se estimaron los ángulos formados entre articulaciones clave, lo que permite entender la postura del cuerpo en diferentes actividades.
- **Inclinación del tronco:** Se midió la inclinación del tronco comparando la posición de los hombros y las caderas, proporcionando información sobre la postura y el equilibrio del individuo.

Entrenamiento del Sistema de Clasificación

Elección del Modelo

Se seleccionó el modelo **Support Vector Machine (SVM)** por su capacidad para clasificar datos en espacios de alta dimensión y su eficacia en problemas de clasificación binaria y multiclase.

Entrenamiento

Los pasos realizados para el entrenamiento del modelo incluyen:

1. **División de los Datos:** Se separaron los datos en conjuntos de entrenamiento y prueba para asegurar una evaluación adecuada del modelo.
2. **Entrenamiento:** Se utilizó el conjunto de entrenamiento para ajustar el modelo SVM, empleando las características extraídas previamente (posiciones de articulaciones, velocidades, ángulos, etc.).

Resultados

El modelo SVM mostró un rendimiento significativo en la clasificación de las actividades. A continuación, se presentan las métricas de evaluación:

	precision	recall	f1-score	support
CaminandoEspalda_01	1.00	1.00	1.00	3
CaminandoEspalda_02	1.00	0.80	0.89	5
CaminandoEspalda_03	1.00	1.00	1.00	2
CaminandoEspalda_04	0.67	1.00	0.80	2
CaminandoFrontal_01	1.00	1.00	1.00	5
CaminandoFrontal_02	1.00	1.00	1.00	3
CaminandoFrontal_03	0.67	0.50	0.57	4
CaminandoFrontal_04	0.86	0.86	0.86	7
Girando_01	1.00	1.00	1.00	1
Girando_03	1.00	1.00	1.00	2
Girando_04	1.00	1.00	1.00	1
Parandose_01	1.00	1.00	1.00	2
Parandose_02	1.00	1.00	1.00	2
Parandose_03	1.00	1.00	1.00	2
Parandose_04	1.00	1.00	1.00	2
Sentandose_01	0.67	1.00	0.80	2
Sentandose_03	0.50	1.00	0.67	1
Sentandose_04	0.00	0.00	0.00	1
accuracy			0.89	47
macro avg	0.85	0.90	0.87	47
weighted avg	0.89	0.89	0.89	47

- **Accuracy:** 0.89
- **Macro Average:** Precision 0.85, Recall 0.90, F1-Score 0.87
- **Weighted Average:** Precision 0.89, Recall 0.89, F1-Score 0.89

Análisis de Resultados

Los resultados muestran un rendimiento sólido del modelo SVM, aunque hay clases con bajo desempeño como "Sentandose_04", donde la precisión y el recall son nulos. Esto podría indicar la necesidad de más datos para esta clase en particular o la dificultad de la tarea de clasificación en esos casos. La mayoría de las clases obtuvieron resultados positivos, lo que sugiere que el modelo está generalizando bien en las actividades comunes.

Conclusiones y Trabajo Futuro

El proyecto ha permitido desarrollar un sistema que clasifica actividades humanas con un alto grado de precisión. Se ha aprendido sobre la importancia del preprocesamiento y la selección de características. Para el futuro, se recomienda:

- Ampliar el conjunto de datos, especialmente para las clases con bajo rendimiento.
- Investigar otras arquitecturas de modelos que podrían mejorar aún más la precisión y el recall.
- Implementar una interfaz gráfica de usuario que permita visualizar la actividad en tiempo real y las medidas posturales.

Referencias

- MediaPipe: [MediaPipe Documentation](#)
- LabelStudio para anotación: [LabelStudio](#)
- CVAT: [CVAT](#)