



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

**Centro Nacional de Investigación
y Desarrollo Tecnológico**

Tesis de Maestría

**Detección de Anomalías en el Comportamiento de
Una Persona**

presentada por

Ing. Gloria Vanessa Urquiza Flores

como requisito para la obtención del grado de
Maestra en Ciencias de la Computación

Director de tesis

Dr. Raúl Pinto Elías

Codirector de tesis

Dr. Gerardo Reyes Salgado

Cuernavaca, Morelos, México. Octubre de 2022.



Centro Nacional de Investigación y Desarrollo Tecnológico
Departamento de Ciencias Computacionales

Cuernavaca, Mor., **29/septiembre/2022**

OFICIO No. DCC/078/2022

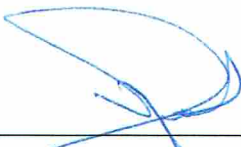
Asunto: Aceptación de documento de tesis
CENIDET-AC-004-M14-OFICIO

DR. CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de la C. Gloria Vannesa Urquiza Flores, con número de control M20CE089, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **“DETECCIÓN DE ANOMALÍAS EN EL COMPORTAMIENTO DE UNA PERSONA”** y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.



Dr. Raúl Pinto Elías
Director de tesis

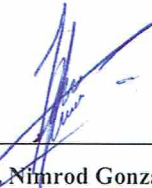


Dr. Gerardo Reyes Salgado
Codirector de Tesis

Dr. José Ruiz Ascencio
Revisor 1



Dra. Andrea Magadán Salazar
Revisor 2



Dr. Nimrod González Franco
Revisor 3

C.c.p. Depto. Servicios Escolares.
Expediente / Estudiante
JGGS/ibm



Interior Internado Palmira S/N, Col. Palmira, C. P. 62490, Cuernavaca, Morelos
Tel. 01 (777) 3627770, ext. 3201, e-mail: dcc@tecnm.mx | cenidet.tecnm.mx





Cuernavaca, Mor.,

05/octubre/2022

No. De Oficio:

SAC/150/2022

Asunto:

Autorización de
impresión de tesis

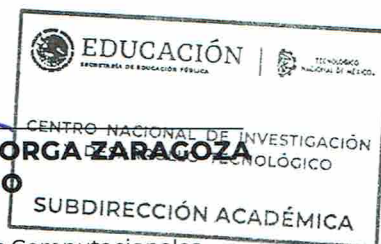
**GLORIA VANNESA URQUIZA FLORES
CANDIDATO AL GRADO DE MAESTRO(A) EN CIENCIAS
DE LA COMPUTACIÓN
P R E S E N T E**

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado "DETECCIÓN DE ANOMALÍAS EN EL COMPORTAMIENTO DE UNA PERSONA", ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE

Excelencia en Educación Tecnológica®
"Educación Tecnológica al Servicio de México"



DR. CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO

C. c. p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/CHG



Dedicatorias

El presente trabajo va dedicado a Dios, guía quien ha estado presente en el caminar de mi vida, bendiciéndome y dándome fuerza para continuar mis metas, sin desfallecer.

Con todo mi cariño se lo dedico a mi madre, por ser una gran mujer. El esfuerzo y las metas alcanzadas, refleja la dedicación, el amor que invierte una madre en sus hijos.

A mi padre, por su amor incondicional, por inculcarme no temer las adversidades porque Dios está conmigo siempre.

A mis hermanas, Karla y Gabriela, por acompañarme en todos mis sueños y metas.

Con cariño, Gloria

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada durante la realización de esta investigación.

Agradezco a todos los investigadores, técnicos y personal administrativo y de apoyo del TecNM / CENIDET; asimismo a mis compañeros de generación, por su compañía y apoyo durante mi formación.

Un especial agradecimiento a todos los académicos del departamento de Ciencias computacionales, que fueron participantes de esta investigación, ya que sin su apoyo y colaboración esta tesis no habría sido posible.

Quiero agradecerle a mi director de tesis, el Dr. Raúl Pinto Elías, sus conocimientos invaluable que me brindó para llevar a cabo y culminar esta investigación.

Al Codirector de esta tesis, el Dr. Gerardo Reyes Salgado, por las valiosas contribuciones a esta investigación y por el tiempo dedicado para revisarlo.

De igual manera agradezco el apoyo y dedicación, al comité revisor, la Dra. Andrea Magadán Salazar, el Dr. José Ruiz Ascencio y el Dr. Nimrod González Franco, por el tiempo prestado y sugerencias para la mejora de este trabajo.

Finalmente, un agradecimiento con mucho cariño, a mi familia por acompañarme en esta etapa de mi crecimiento académico.

Resumen

El reconocimiento de actividad humana HAR (Human Activity Recognition) es un área de investigación activa, el interés por desarrollar esta área se debe a su amplia variedad de aplicaciones en el mundo real como, en la videovigilancia, seguridad, deporte, atención médica, vida asistida y monitoreo continuo en el hogar.

El HAR es un tema relacionado con el análisis del comportamiento humano con gran trayectoria de investigación, enfocado al comportamiento individual de un ser humano, analizando las actividades diarias de una persona, para reconocer patrones de actividad de un individuo monitoreado, con el objetivo de determinar la existencia de cambios en el diario vivir de un ser humano.

En el presente trabajo se propone el desarrollo de un modelo multimodal para monitorear las actividades de una persona en cuatro habitaciones de un hogar, utilizando cámaras RGB, sensores acelerómetro y giroscopio, para transferir la información obtenida de este monitoreo a una aplicación de escritorio.

Mediante la observación de lo que suele hacer una persona durante una rutina en un periodo establecido, el modelo podrá utilizar su conocimiento para monitorear la rutina actual de una persona y detectar si existe alguna desviación de los patrones de actividad.

Para evaluar la solución propuesta se recabó el monitoreo de una rutina diaria “normal” de las actividades diarias que una persona suele realizar, de acuerdo a sus necesidades de estilo vida. Posteriormente, se monitorea la rutina actual de una sola persona, para detectar desviaciones en comparación a lo que suele hacer esa persona.

Como resultado se obtuvo que el modelo es capaz de detectar anomalías en el comportamiento de una persona, detectando los siguientes casos en el comportamiento monitoreado: anormalidad en el día, anormalidad en el periodo, patrón anormal, patrón normal, patrón de comportamiento anormal y patrón inhabitual.

Abstract

Human Activity Recognition HAR is an active research area, the interest in developing this area is due to its wide variety of real-world applications such as video surveillance, security, sports, healthcare, life assisted and continuous monitoring at home.

HAR is a subject related to the analysis of human behavior with a long history of research, focused on the individual behavior of a human being, monitoring the daily activities of a person, to recognize activity patterns of a monitored individual, with the aim of determining the existence of changes in the daily life of a human being.

In the present work, the development of a multimodal model is proposed to monitor the activities of a person in four rooms of a home, using RGB cameras, accelerometer and gyroscope sensors, to transfer the information obtained from this monitoring to a desktop application.

By observing what a person typically does during a routine over a set period, the model will be able to use its knowledge to monitor a person's current routine and detect if there is any deviation from activity patterns.

To evaluate the proposed solution, the monitoring of a "normal" daily routine of the daily activities that a person usually performs, according to their needs and lifestyle, was collected. Subsequently, the current routine of a person is monitored, to detect deviations compared to what a person usually does.

As a result, it was obtained that the model is capable of detecting anomalies in the behavior of a person. Detecting the following cases in the monitored behavior: abnormality in the day, abnormality in the period, abnormal pattern, normal pattern, abnormal behavior pattern and unusual pattern.

Índice

Lista de figuras	IV
Lista de tablas	VI
Glosario	VIII
Acrónimos	IX
CAPÍTULO 1 INTRODUCCIÓN.....	1
1.1 Descripción del problema	2
1.1.1 Delimitación del problema específico	2
1.1.2 Complejidad del problema	2
1.2 Marco Conceptual	3
1.2.1 Visión por computadora	3
1.2.2 Estimación de la pose humana	3
1.2.3 Detección de objetos	4
1.2.4 Sensores	4
1.2.5 Monitoreo de la rutina	4
1.2.6 Detección de anomalías	5
1.2.7 Máquina de estado finito	5
1.3 Organización de la tesis	6
Capítulo 2 ESTADO DEL ARTE	7
2.1 Reconocimiento de actividades humanas	7
2.2 Visión por computadora	8
2.3 Estimación de la pose humana	9
2.4 Modelado del comportamiento	10
2.5 Reconocimiento de anomalía.....	11
2.6 Sensores.....	12
2.7 Antecedentes	13
2.8 Discusión	14
CAPÍTULO 3 PROPUESTA DE SOLUCIÓN.....	15
3.1 Planteamiento de solución	15
3.1.1 Objetivo	15
3.1.2 Alcances y limitaciones	16
3.1.3 Justificación y beneficios	17
3.2 Modelo conceptual de la solución.....	17
3.3 Descripción de los elementos del modelo conceptual	18

3.4	Variables / criterios de control.....	20
CAPÍTULO 4 ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DEL SISTEMA		21
4.1	Arquitectura del sistema.....	21
4.2	Etapa de localización de una persona.....	24
4.2.1	Sensor RFID.....	25
4.2.2	Pulsadores eléctricos	26
4.2.3	App móvil.....	27
4.2.4	Adquisición de video	28
4.3	Etapa de estimación de la pose humana.....	28
4.3.1	Open Pose.....	29
4.3.2	MediaPipe Pose.....	30
4.4	Etapa de reconocimiento de objetos	32
4.4.1	Yolov3.....	33
4.5	Etapa de recolección de datos de sensores	35
4.6	Etapa de Reconocimiento de actividad.....	37
4.6.1	Medida de proximidad.....	39
4.6.2	Secuencia de actividades.....	40
4.7	Etapa de modelado de rutina	40
4.7.1	Grafo dirigido	40
4.8	Etapa de reconocimiento de anomalía.....	41
4.8.1	Máquina de estado finito	42
Capítulo 5 VALIDACIÓN Y EXPERIMENTACIÓN.....		43
5.1	Ambiente de pruebas	43
5.2	Banco de datos	44
5.3	Experimentación.....	44
5.4	Etapa de localización de una persona.....	44
5.5	Etapa de estimación de la pose humana.....	45
5.6	Etapa de reconocimiento de objetos	52
5.7	Etapa de recolección de datos de sensores	57
5.8	Etapa de reconocimiento de actividad.....	64
5.9	Etapa de modelado de la rutina	69
5.10	Etapa de reconocimiento de anomalía.....	73
5.11	Plan de pruebas	78
5.11.1	Entrenamiento del modelo.....	78
5.11.2	Validación del modelo.....	82

5.12	Discusión	
Capítulo 6	ANÁLISIS DE RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO	91
6.1	Análisis de resultados.....	91
6.2	Conclusión.....	92
6.3	Aportaciones	93
6.4	Trabajos futuros.....	93
	Referencias.....	94
	Anexo A. Open Pose.....	99
	Anexo B. Artículo	101

Lista de figuras

Figura 3.1. Diagrama de componentes	17
Figura 4.1. Arquitectura del sistema.....	23
Figura 4.2. Diagrama de distribución de cámaras RGB.....	24
Figura 4.3. Lector RFID	25
Figura 4.4. Tarjetas (tags).....	26
Figura 4.5. Lector RFID y tarjetas.....	26
Figura 4.6. Pulsadores eléctricos	26
Figura 4.7. Diseño de control de pulsadores eléctricos	27
Figura 4.8. Interfaz gráfica de aplicación móvil.....	28
Figura 4.9. Proceso del método propuesto en Open Pose	29
Figura 4.10. Codificar un vector 2D Open Pose	30
Figura 4.11. Resultado de estimación de la pose Open Pose	30
Figura 4.12. Hombre de Vitruvio.	31
Figura 4.13. Puntos de referencia de la pose de MediaPipe Pose	32
Figura 4.14. Estructura de la red Darknet-53	33
Figura 4.15. Habitación Cocina.....	34
Figura 4.16. Habitación Dormitorio	34
Figura 4.17. Habitación Sala	35
Figura 4.18. Habitación Dormitorio	35
Figura 4.19. Interfaz gráfica de aplicación móvil.....	36
Figura 4.20. Representación de cálculo de medida de proximidad.....	39
Figura 5.1. Caso rostro no visible: usar microondas	46
Figura 5.2. Caso rostro no visible: usar microondas	46
Figura 5.3. Caso rostro no visible: usar microondas	46
Figura 5.4. Caso rostro no visible: leer periódico.....	47
Figura 5.5. Caso rostro no visible: leer periódico.....	47
Figura 5.6. Caso rostro no visible: leer periódico.....	47
Figura 5.7. Caso rostro no visible: salir de casa	47
Figura 5.8. Caso rostro no visible: salir de casa	47
Figura 5.9. Caso rostro no visible: salir de casa	48
Figura 5.10. Caso rostro no visible: salir de casa	48
Figura 5.11. Caso rostro no visible: tomar fruta.....	48
Figura 5.12. Caso rostro no visible: tomar fruta.....	48
Figura 5.13. Caso rostro no visible: tomar fruta.....	49
Figura 5.14. Caso obstrucción de la postura: preparar alimento	49
Figura 5.15. Caso obstrucción de la postura: preparar alimento	49
Figura 5.16. Caso obstrucción de la postura: preparar alimento	49
Figura 5.17. Caso obstrucción de la postura: comer.....	49
Figura 5.18. Caso obstrucción de la postura: comer.....	50
Figura 5.19. Caso obstrucción de la postura: comer.....	50

Figura 5.20. Caso falsa estimación de la postura	50
Figura 5.21. Caso falsa estimación de la postura	50
Figura 5.22. Caso falsa estimación de la postura	51
Figura 5.23. Caso falsa estimación de la postura	51
Figura 5.24. Detección de objetos: Cocina.....	53
Figura 5.25. Detección de objetos: Comedor	53
Figura 5.26. Detección de objetos: Sala	54
Figura 5.27. Detección de objetos: Dormitorio	55
Figura 5.28. Etiquetado manual de objetos: Cocina.....	55
Figura 5.29. Etiquetado manual de objetos: Comedor	56
Figura 5.30. Etiquetado manual de objetos: Sala	56
Figura 5.31. Etiquetado manual de objetos: Dormitorio	56
Figura 5.32. Matriz de confusión: WISDM.....	59
Figura 5.33. Matriz de confusión: conjunto propio	61
Figura 5.34. Representación de espectro de ruido en acelerómetro	62
Figura 5.35. Representación de filtrado de la señal del acelerómetro	62
Figura 5.36. Representación de espectro de ruido en giroscopio	63
Figura 5.37. Representación de filtrado de la señal del giroscopio.....	63
Figura 5.38. Cálculo de medida de proximidad	64
Figura 5.39. Secuencia de actividades.....	68
Figura 5.40. Grafo: Cocina.....	69
Figura 5.41. Grafo: Comedor	70
Figura 5.42. Grafo: Sala	70
Figura 5.43. Grafo: Dormitorio	71
Figura 5.44. Grafo: General.....	72
Figura 5.45. Resultado de procesar un video.....	80
Figura 5.46. Resultado de procesar un video.....	80
Figura 5.47. Visualizador de imágenes reconocidas por el modelo	81
Figura 5.48. Resultado prueba 1	83
Figura 5.49. Resultado Prueba 1.....	83
Figura 5.50. Resultado de prueba 2 en formato txt	84
Figura 5.51. Resultado de prueba 2 en interfaz gráfica	85
Figura 5.52. Resultado de prueba 3 formato txt	86
Figura 5.53. Resultado de prueba 3 en interfaz gráfica	86
Figura 5.54. Resultado de prueba 4 en formato txt	87
Figura 5.55. Resultado de prueba 5 en interfaz gráfica.....	88

Lista de tablas

Tabla 4.1. Preprocesamiento de los datos del sensor	36
Tabla 4.2. Lista de Actividades Reconocidas.....	38
Tabla 4.3. Descripción de Grafo dirigido.....	41
Tabla 4.4. Definición formal de máquina de estado finito	42
Tabla 5.1. Características de banco de datos	44
Tabla 5.2. Resultados de reconocimiento de actividad: Cocina.....	51
Tabla 5.3. Resultados de reconocimiento de actividad: Comedor	51
Tabla 5.4. Resultados de reconocimiento de actividad: Sala	52
Tabla 5.5. Resultados de reconocimiento de actividad: Dormitorio	52
Tabla 5.6. Resultados de detección de objeto: Cocina	53
Tabla 5.7. Resultados de detección de objeto: Comedor.....	53
Tabla 5.8. Resultados de detección de objeto: Sala.....	54
Tabla 5.9. Resultados de detección de objeto: Dormitorio.....	54
Tabla 5.10. Etiquetado manual de objetos.....	55
Tabla 5.11. Descripción banco de datos WISDM	57
Tabla 5.12. Modelo de red neuronal profunda	58
Tabla 5.13. Resultados de WISDM.....	59
Tabla 5.14. Descripción de banco de datos propio.....	60
Tabla 5.15. Resultados de conjunto de datos propio	60
Tabla 5.16. Reconocimiento de actividades	65
Tabla 5.17. Tabla de transiciones grafo: Cocina	69
Tabla 5.18. Tabla de transiciones grafo: Comedor.....	70
Tabla 5.19. Tabla de transiciones grafo: Sala.....	71
Tabla 5.20. Tabla de transiciones grafo: Dormitorio.....	71
Tabla 5.21. Tabla de transiciones grafo: General	72
Tabla 5.22. Definición formal de máquina de estado finito: Cocina.....	73
Tabla 5.23. Tabla de transiciones máquina de estado finito: Cocina	74
Tabla 5.24. Definición formal de máquina de estado finito: Comedor	74
Tabla 5.25. Tabla de transiciones máquina de estado finito: Comedor.....	74
Tabla 5.26. Definición formal de máquina de estado finito: Sala	74
Tabla 5.27. Tabla de transiciones máquina de estado finito: Sala.....	75
Tabla 5.28. Definición formal de máquina de estado finito: Dormitorio	75
Tabla 5.29. Tabla de transiciones máquina de estado finito: Dormitorio.....	75
Tabla 5.30. Definición formal de máquina de estado finito: General	76
Tabla 5.31. Tabla de transiciones máquina de estado finito: General	77
Tabla 5.32. Entrenamiento del modelo: Rutina normal.....	79
Tabla 5.33. Entrenamiento: Datos resultado de procesar un video	79
Tabla 5.34. Validación: Datos resultado de procesar un video.	82
Tabla 5.35. Prueba 2: Rutina de actividades	84
Tabla 5.36. Prueba 3: Rutina de actividades	85

Tabla 5.37. Prueba 4: Rutina de actividades	87
Tabla 5.38. Prueba 5: Rutina de actividades	88
Tabla 6.1. Definición de eventos clasificados por el modelo.....	91
Tabla A.1. Resultados de Open Pose.....	99

Glosario

Palabra	Definición
Android Studio	Entorno de desarrollo integrado oficial para la plataforma Android
BlazePose	Arquitectura de red neuronal convolucional ligera para la estimación de la pose humana
CENIDET	Centro Nacional de Investigación y Desarrollo Tecnológico
2D	2 dimensiones
3D	3 dimensiones
Darknet-19	Red neuronal convolucional que se utiliza en YOLOv2
Distancia <i>Manhattan</i>	Longitud entre dos puntos
<i>eWatch</i>	Dispositivo reloj inteligente
Hz	Hercio unidad de medida de la frecuencia
<i>Human Pose Estimation</i>	Estimación de la pose humana
Kinect	Controlador de juego y entretenimiento desarrollado por Microsoft©
<i>Machine Learning</i>	Aprendizaje automático
<i>MediaPipe Pose</i>	Solución de aprendizaje automático para el seguimiento de la postura del cuerpo
<i>Open Pose</i>	Sistema en tiempo real para la detección de pose 2D de varias personas
ResNet-101	Red neuronal convolucional que tiene 101 capas de profundidad
RetinaNet	Modelos de detección de objetos
<i>Spyder</i>	Entorno de desarrollo integrado para programación científica en el lenguaje Python
UTD-HAD	Conjunto de datos público

Acrónimos

Acrónimo	Significado
AVD	Actividades de la vida diaria
BOW	<i>Bag Of Word</i> (Bolsa de palabra)
CCTV	Circuito Cerrado de Televisión
CNN	<i>Convolutional Neural Networks</i> (Red neuronal convolucional)
FPS	<i>Frames per second</i> (Frames por segundo)
FSM	<i>Finite State Machine</i> (Máquina de estados finitos)
GPS	<i>Global Positioning System</i> (Sistema de posicionamiento global)
HAR	<i>Human Activity Recognition</i> (Reconocimiento de actividad humana)
HD	<i>High-definition</i> (Alta definición)
HOF	Histograma de descriptores de flujo
HOG	Histograma de gradiente orientado
HMM	Modelos Ocultos de Markov
IDE	<i>Integrated Development Environment</i> (Entorno de desarrollo integrado)
KNN	<i>K-Nearest Neighbors</i> (K-vecino más cercano)
LSTM	<i>Long Short Term Memory</i> (Red de memoria a largo plazo)
MEI	Imagen de energía de movimiento
MHI	Imagen historial de movimiento
NN	<i>Neural networks</i> (Redes neuronales)
PAF	<i>Part Affinity Fields</i> (Campos de afinidad parcial)
RAM	<i>Random Access Memory</i> (Memoria de acceso aleatorio)
RGB	<i>Red, Green, Blue</i> (Rojo, Verde, Azul)
RFID	<i>Radio Frequency Identification</i> (Identificación por radiofrecuencia)
RVM	Máquinas de Vectores de Relevancia
STIP	<i>Space-Time Interesting Point</i> (Punto de interés espacio-temporal)
SVM	Máquinas de Soporte Vectorial
YOLO	<i>You only look once</i> Sistema de detección de objetos

CAPÍTULO 1

INTRODUCCIÓN

El reconocimiento automático de actividades físicas es conocido como Reconocimiento de Actividad Humana HAR, tiene por objetivo el estudio y modelado de los movimientos que provienen del cuerpo humano al realizar una actividad (J. Thummala, et al., 2020).

Actualmente, el reconocimiento de actividades, está desarrollado para aplicaciones como para atención médica (M. Skubic, et al., 2015), vida asistida (Mettel, et al., 2019), vigilancia (Hossen & Tuli, 2016), seguridad, deporte, entre otras. Se utilizan diferentes metodologías para monitorear a los seres humanos y sus actividades diarias.

Los sensores más comunes para captar movimientos humanos son las cámaras de video, sensores de profundidad, giroscopios, acelerómetros, entre otros. Aunque cada sensor tiene ventajas significativas, utilizar un solo tipo de sensor para monitorear actividades ha mostrado ser insuficiente para caracterizar una amplia variedad de actividades humanas.

La combinación de diferentes fuentes de información es una opción para mejorar la caracterización de las actividades diarias aprovechando las ventajas de cada sensor. En esta tesis, se propone el desarrollo de un sistema multimodal utilizando cámaras RGB, además de los sensores acelerómetro y giroscopio.

El reconocimiento de actividades humanas está relacionado con el comportamiento humano, puesto que es un tema con trayectoria de investigación activa mayormente enfocado al análisis individual de una persona (Yin, Yang, & Pan, 2008), para detectar desviaciones en la rutina habitual, centrándose en el procesamiento de secuencias de videos que proponen aplicaciones para ofrecer una mejor calidad de vida, monitoreando el diario vivir de las personas.

1.1 Descripción del problema

La detección automática del comportamiento humano de una persona monitoreada en secuencia de videos es un problema de Visión por Computador y *Machine Learning* que tiene retos que todavía los investigadores deben superar, (Alshurafa, et al., 2017).

Las diferentes técnicas para la detección de anomalías tienen ventajas y desventajas, es importante seleccionar cuál es la más adecuada para un determinado problema de detección de anomalías (Skubic, Guevara, et al., 2015), (Pavel, et al., 2015).

La definición exacta de anomalía puede cambiar de acuerdo al dominio y área de aplicación. Por ejemplo, lo que en una determinada situación es normal en otra puede considerarse anormal. Una pelea en la calle es una actividad anormal, mientras que una pelea en un ring de boxeo es normal.

1.1.1 Delimitación del problema específico

El problema a investigar es analizar el comportamiento humano, reconociendo patrones anómalos en una rutina cotidiana realizada por una persona, en comparación con el comportamiento aprendido de la persona monitoreada.

Es decir, dada una rutina realizada por una persona dentro de un hogar, se pretende que el modelo sea capaz de reconocer una desviación en el comportamiento cotidiano de la persona.

1.1.2 Complejidad del problema

La complejidad del problema se estructura de la siguiente forma, identificando lo que el sistema HAR debe enfrentar:

1. Captura de datos reales y suficientes para establecer relaciones en los cambios del comportamiento del ser humano estudiado.
2. Un aspecto crítico para un método de detección de anomalías es determinar cuáles son normales y cuáles son anormales.
3. Variabilidad individual de los seres humanos en los patrones de comportamiento.
4. Cambios de comportamiento que no son cuantificables y son difíciles de observar.
5. Desafíos técnicos como la duración de la batería del dispositivo portátil, aceptación del usuario, cumplir con el monitoreo a largo plazo.
6. Crear un patrón de comportamiento con la rutina de un humano implementando un modelo de *Machine Learning*.

1.2 Marco Conceptual

El reconocimiento de actividades humanas puede ser realizado por una variedad de métodos y técnicas utilizadas en investigaciones. A continuación, se resumen las técnicas utilizadas que forman una base teórica en la cual se fundamenta esta investigación. En esta sección, se definen los conceptos teóricos como, la estimación de la pose, el monitoreo de la rutina, la detección de anomalías.

1.2.1 Visión por computadora

Para el ser humano, el reconocimiento de actividad se considera una habilidad totalmente natural. Tratar de emular este tipo de habilidad en sistemas de detección automático se considera un gran desafío y ha motivado a investigadores en el área a proponer diferentes tipos de soluciones, un enfoque común es el uso de técnicas de procesamiento digital de imágenes y visión por computador.

Dentro de sus campos de estudio, la visión artificial se encarga de analizar, procesar y sobre todo entender la información contenida en imágenes del mundo real a través de diferentes técnicas. Estas técnicas utilizan procedimientos de segmentación de imagen y extracción de características.

Una técnica popular es la sustracción de fondo de la escena, este método permite separar el fondo de los objetos en movimiento. En la extracción de características generalmente se utilizan descriptores tales como la forma, silueta, colores, movimientos del cuerpo y poses.

1.2.2 Estimación de la pose humana

La visión artificial tiene amplias líneas de investigación. La dificultad de entender con exactitud imágenes digitales del mundo real hace que muchas de estas líneas tengan una gran cantidad de retos aún sin resolver, una de estas líneas de investigación es la estimación de la pose humana.

La estimación de la pose humana (*Human Pose Estimation*) es el proceso de predecir las posiciones 2D o 3D de cada una de las partes del cuerpo de una o varias personas que se encuentren presentes en imágenes o videos (Cao, Hidalgo, et al., 2021).

Los métodos de estimación de la pose humana pueden procesar imágenes o videos y transforman la información numérica y simbólica para entender el contenido de la imagen, obteniendo como resultado de la variedad de modelos para la estimación humana la configuración espacial de las partes del cuerpo humano en imágenes o videos.

1.2.3 Detección de objetos

La visión es uno de los mecanismos de percepción más importantes que tiene el ser humano, es utilizada para desenvolverse eficientemente dentro del ambiente que nos rodea y para detectar los objetos de interés a través de su forma, color, relieve y dimensiones.

La visión por computadora es la capacidad de la máquina para ver el mundo que lo rodea, para deducir estructuras y propiedades a partir de imágenes. Una línea de investigación activa en el área es la implantación de habilidades en una máquina para detectar y determinar la identidad de los objetos, (Mahendru & Dubey, 2021).

Uno de los desafíos importantes es que no existe algún algoritmo eficaz que reconozca todo tipo de objetos en cualquier ambiente y en el tiempo en que el sentido sensorial de la vista lo realiza, (Redmon, et al., 2018).

1.2.4 Sensores

En el área de reconocimiento de actividades humanas existe una línea de investigación activa por realizar la detección automática (Fu, Damer, et al., 2020), reconocimiento y análisis de acciones humanas a partir de datos obtenidos de sensores.

La detección de acciones simples y complejas implica determinar la presencia de la acción de interés dentro de un flujo de datos. Un sensor es un dispositivo que permite la transducción de cualquier magnitud física o química en una señal eléctrica aprovechando las propiedades eléctricas de la materia (Mahmud, et al., 2020).

Usualmente, los dispositivos utilizados para el registro de actividades físicas son sensores de movimiento, estos son dispositivos eléctricos equipados para obtener una respuesta eléctrica cuando se ejecuta una acción física.

Los sensores más comunes para el monitoreo de actividades humanas son los acelerómetros, giroscopios, unidades de masa inercial (Bruno, et al., 2013), entre otros. Con el avance tecnológico se tiene al alcance una variedad de sensores en dispositivos móviles como reloj o teléfono inteligente.

1.2.5 Monitoreo de la rutina

En el área de reconocimiento de actividad humana, ha habido un mayor enfoque en el desarrollo de tecnología que permita la evaluación continua de los cambios basados en el comportamiento y patrones de actividad de un individuo monitoreado, con el objetivo de determinar la existencia de cambios en el diario vivir de un ser humano.

En este tipo de sistemas se utilizan sensores integrados en el entorno, para recopilar patrones de comportamiento y actividad durante una rutina. Una rutina está asociada a realizar diferentes actividades diarias de manera periódica por un ser humano.

Existe un gran desafío en esta área, que es la captura de datos reales y suficientes para establecer relaciones en los cambios del comportamiento de una persona, además de identificar los mejores parámetros para realizar un seguimiento en los patrones de comportamiento.

1.2.6 Detección de anomalías

La detección de anomalías en el comportamiento de una persona, es un problema estudiado en diferentes áreas de investigación y sus dominios de aplicación. La palabra anomalía puede tener diferentes significados dependiendo del campo de investigación en que se utilice. Algunas de las áreas donde se utiliza la detección de anomalías es en la medicina, vida asistida, monitoreo de adultos mayores, entre otros.

El modelo propuesto en esta investigación realiza el reconocimiento de anomalías en el comportamiento de una persona. Primero el modelo es entrenado con datos de una rutina “normal” para una persona. Esta rutina “normal” contiene actividades de la vida diaria que una persona hace de acuerdo con días de la semana, período, estilo de vida y necesidades.

El reconocimiento de un comportamiento “anormal” se logra cuando el modelo detecta una desviación en el comportamiento “normal” con el que fue entrenado el modelo, entonces para el modelo de esta investigación la detección de una anomalía será identificada por la diferencia que existe entre el comportamiento aprendido por el modelo y el comportamiento actual.

1.2.7 Máquina de estado finito

Una máquina de estados finitos se puede definir como un modelo computacional que acepta símbolos de entrada, genera símbolos de salida y cambia su estado interno de acuerdo con algunas reglas predefinidas (Yuan & Xu, 2011). El cálculo comienza en el estado inicial con una cadena de entrada. Cambia a nuevos estados dependiendo de la función de transición. Formalmente, un autómata finito es una 5-tupla $(Q, \Sigma, q_0, \delta, F)$ (Chakraborty, 2003) donde:

- Q : Conjunto finito de estados.
- Σ : Alfabeto finito.
- $q_0 \in Q$: Estado inicial.
- $\delta: Q \times \Sigma \rightarrow Q$: Función de transición.
- $F \subseteq Q$: Conjunto de estados finales o de aceptación.

1.3 Organización de la tesis

El capítulo 2 muestra el estudio del estado del arte, dando a conocer las diferentes técnicas para el reconocimiento de actividades humanas, además de investigaciones enfocadas al reconocimiento de anomalías en el comportamiento humano.

En el capítulo 3 se detallan el objetivo, los alcances y las limitaciones, así como la justificación de esta investigación. También se presenta el planteamiento y el modelado conceptual de la solución.

En el capítulo 4 se presenta el análisis y el diseño del sistema y además de que se describen las diferentes etapas de implementación.

En el capítulo 5 se detalla la validación y experimentación del modelo, además de un plan de pruebas realizado para mostrar el cumplimiento de los objetivos planteados.

En el capítulo 6 se presentan las conclusiones de este trabajo de investigación y se mencionan los trabajos futuros que surgieron en el desarrollo de la investigación.

Capítulo 2

Estado del Arte

En este capítulo se hace una revisión del estado del arte del problema relacionado con el reconocimiento de actividades humanas y la detección de anomalías en el comportamiento humano.

En la primera sección del estado del arte 2.1 se estudian diferentes propuestas de sistemas que realizan reconocimiento de actividades humanas, seguido de la sección 2.2 donde se puede encontrar sistemas que utilizan solamente técnicas de visión por computadora para el reconocimiento de actividades; en la sección 2.3 se destacan investigaciones que han implementado la estimación de la pose humana, mientras que en la sección 2.4 se encuentran las investigaciones relacionadas al monitoreo continuo de las actividades humanas para realizar el modelado de comportamiento, en la sección 2.5 se presentan investigaciones que se enfocan el reconocimiento de anomalías en el comportamiento de una persona, en la sección 2.6 se mencionan investigaciones que implementan el uso de sensores portátiles.

En la sección 2.7, se encuentran los trabajos antecedentes a esta investigación y finalmente, se presenta un análisis de lo descrito en la sección 2.8.

2.1 Reconocimiento de actividades humanas

El reconocimiento y clasificación de actividades humanas busca, a partir de cierta fuente de datos, detectar la actividad que está siendo realizada por determinada persona, a través del análisis de un conjunto de datos recolectados. Mediante el desarrollo y surgimiento de nuevos dispositivos y tecnologías, ha sido posible la creación de diversas aplicaciones en esta área de investigación.

En el artículo (Bobick & Davis, 2001), se presenta un enfoque que contrasta esfuerzos para recuperar la forma humana a partir de la imagen, considerando el reconocimiento de acción dentro de un marco basado en el movimiento. Se utiliza una plantilla de imagen vectorial, generando una imagen de energía de movimiento (MEI) y una imagen historial de movimiento (MHI). Para la clasificación de la actividad, las plantillas se comparan con

modelos almacenados de vistas de acciones conocidas, logrando el reconocimiento de 18 ejercicios aeróbicos.

En el artículo (Babiker, et al., 2017), se detecta el cuerpo humano a través de la resta de fondo 2D para un sistema de vigilancia inteligente, se propone la estimación de un cuadro delimitador para rastrear el cuerpo humano y se realiza el reconocimiento del tipo de movimiento y actividad, se logra reconocer las actividades: caminar, sentarse, boxear, agitar las manos y colocar.

En el artículo (Zhen-Yu He, et al., 2008), se presenta un modelo autorregresivo de series de tiempo para reconocer la actividad humana a partir de los datos de un acelerómetro triaxial ubicado en el bolsillo del pantalón. La clasificación de las actividades humanas se realiza con *Support Vector Machine* (SVM). Los resultados de reconocimiento promedio para cuatro actividades (correr, quieto, saltar y caminar) usando las características propuestas son de 92.25%, que son mejores que usando características tradicionales de dominios de tiempo de uso frecuente: media, desviación estándar, energía y correlación de datos de aceleración.

Un sistema de reconocimiento de actividad diaria (Zhu & Sheng, 2009) en vida asistida por un robot, infiere situaciones humanas a partir de los datos de movimiento y los signos vitales del humano. El sistema utilizó dos sensores inerciales, conectados en un pie y la cintura del sujeto. Se realiza la recopilación de la aceleración 3D y la velocidad angular, las cuales se fusionan y se clasifican mediante el algoritmo basado en la fusión de múltiples sensores que combina redes neuronales (NN) y modelos ocultos de Markov (HMM); el resultado de precisión es de 90%.

2.2 Visión por computadora

El reconocimiento de actividades humanas HAR ha sido campo de investigación tradicionalmente a través del área de visión artificial. Muchas de las investigaciones desarrolladas enfocadas a sistemas de monitoreo y vigilancia (Hossen & Tuli, 2016), se basan en la utilización de imágenes o video como fuente de datos. Principalmente, buscan predecir actividades o comportamientos de una persona con base en videos capturados de actividades realizadas.

En un sistema de detección de caídas en adultos mayores (Thummala & Pumrin, 2020), se utilizan cámaras RGB ubicadas en la pared de un entorno interior, de las que se obtiene información simple de movimiento y forma 2D, se utiliza un método de resta de fondo se calcula una elipse aproximada que proporciona información sobre la forma y orientación de la persona. Para discriminar una caída, se identifica si ha ocurrido un gran movimiento y se calcula el cambio en la forma humana, si se detecta una elipse inmóvil, entonces, la caída es confirmada por el sistema.

En el artículo (Oikonomopoulos, et al., 2006) se detectan características espacio-temporales en secuencias de imágenes. Se realiza la representación de una acción humana como un conjunto de puntos espacio-temporales. Se mostró el método propuesto usando secuencias de imágenes reales, con la implementación de ejercicios aeróbicos. Se evalúa un clasificador simple de vecino más cercano K (KNN) y uno basado en máquinas de vectores de relevancia (RVM).

En el artículo (Mettel, et al., 2019), el sistema se basa en el seguimiento del esqueleto y las imágenes de profundidad generadas con Microsoft Kinect®. El seguimiento del esqueleto se utiliza para la localización del usuario y para estimar la trayectoria de movimiento. La detección estática verifica si una persona está acostada en el suelo. La detección dinámica comprueba si una persona está cayendo al suelo en ese momento. Si la persona está tendida en el suelo y previamente se ha caído al suelo, se detecta una caída.

Los resultados experimentales de (Mahjoub & Atri, 2016), muestran que el método propuesto, donde se representa la secuencia de acción en video mediante una bolsa espacio-temporal de características basadas en el punto interés espacio-tiempo ($STIP$), además el histograma de gradiente orientado (HOG) y el histograma de descriptores de flujo óptico (HOF). Como clasificador utilizan el SVM que realiza la comparación de características para la clasificación de 27 acciones, se obtienen resultados prometedores y una precisión más alta, se utilizó el conjunto de datos público $UTD-HAD$.

2.3 Estimación de la pose humana

La estimación de la postura humana es una de las áreas desafiantes, pero ampliamente investigadas (Cao, Hidalgo, et al., 2021), puesto que su implementación es necesaria en múltiples tareas de visión artificial que incluyen el reconocimiento de actividad humana, monitoreo humano, videojuegos, lenguaje de señas y videovigilancia, entre otras. Se requieren imágenes y videos para cada aplicación usando una cámara RGB estándar, sin ningún dispositivo externo.

Un sistema de monitoreo de gimnasio en (Ganesh, Idgahi, et al., 2020) utilizando una cámara para registrar actividades como flexiones, sentadillas, planchas, estocada hacia adelante y abdominales. Se realiza la extracción de características basadas en puntos del esqueleto usando el modelo *Open Pose* (Cao, Hidalgo, et al., 2021) para cada fotograma se generó un total de 18 puntos del esqueleto. Los puntos generados se almacenan y se introducen a los algoritmos de clasificación, árbol de decisión, máquina de vectores de soporte (SVM), clasificador de vecino más cercano K (KNN) y el clasificador *Random Forest*.

En la investigación (Hu, et al., 2020), se implementa la fusión de múltiples sensores para monitorear actividades de la vida diaria. Se modelan actividades realizadas por adultos mayores portando relojes inteligentes, combinando equipos de escáneres láser y cámaras. Se utiliza la librería *Open Pose* (Cao, Hidalgo, et al., 2021) para extraer la estructura del esqueleto humano, mediante el uso de transferencia de aprendizaje de imágenes RGB a

imágenes de profundidad, obteniendo la estructura del esqueleto humano con 25 puntos clave.

El artículo (Amrutha, et al., 2021) presenta un enfoque en tiempo real para la detección y el reconocimiento de lenguaje de señas en videos utilizando el método de estimación de pose holística de *MediaPipe* (Bazarevsky, et al., 2020). Este marco holístico detecta los movimientos de múltiples modalidades: expresión facial, gesto de la mano y postura del cuerpo, para el modelo de reconocimiento de lenguaje de señas.

2.4 Modelado del comportamiento

El modelado de comportamiento es una línea de investigación con una amplia variedad de aplicaciones en vida asistida, medicina, monitoreo de adultos mayores. Son sistemas que convierten datos extraídos de sistemas multimodales, en conocimiento útil que permite inferencias conductuales y predicciones. Para el modelado del comportamiento humano en aplicaciones, es necesario adquirir el reconocimiento de actividades humanas.

(Joon-Ho Lim et al., 2008) Proponen un sistema de reconocimiento de actividad diaria para monitorear el comportamiento de una persona mayor, obtienen datos de cinco sensores de presión en el interior del hogar instalados en muebles y pisos; Las actividades diarias son comida, sueño, excreción, salir y descansar. Para evaluar el desempeño del sistema se usarán datos reales de 24 horas de monitoreo en el hogar. El sistema propuesto permite monitorear la vida diaria de una persona mayor que se pretende que viva sola, que permita una vida saludable continua, brindando una señal de advertencia cuando se detecte que realice actividades anormales.

En el artículo (van Kasteren, et al., 2008), se realiza el reconocimiento de siete actividades (salir de casa, ir al baño, ducharse, dormir, preparar el desayuno, preparar la cena y preparar una bebida), el sistema desarrollado monitorea el número de instancias y el porcentaje de tiempo en que ocurren actividades en el conjunto de datos. A través de la detección continua de las rutinas diarias de un sujeto, se puede proporcionar atención personalizada y asistencia eficiente. Esto es especialmente interesante para varios escenarios comunes como la rehabilitación física y mental o el cuidado de pacientes ancianos y pacientes con deterioro cognitivo.

En el artículo (Helaoui, et al., 2011) se enfocan en una serie de observaciones de sensores al realizar el reconocimiento de actividades para detectar y evitar situaciones de emergencia cuando de manera continua se evalúa la conducta del humano estudiado. Con el enfoque en la multitarea se refiere a efectuar más de una actividad de manera simultánea, como, por ejemplo, cocinar y lavar los trastes al mismo tiempo, característica que se presume en las rutinas diarias del mundo real. Se hace uso de la lógica de Markov para reconocer actividades de la vida diaria AVD intercaladas.

Un sistema de monitoreo continuo y discreto en el hogar (Skubic, et al., 2015), utiliza sensores integrados en el entorno para recopilar patrones de comportamiento y actividad a fin de detectar cambios en la salud, enfocándose en los adultos mayores. El sistema de

monitoreo de actividades diarias se lleva a cabo con 11 sensores de movimiento, un sensor de cama y un sensor de temperatura para capturar la actividad de la estufa y el horno, que están relacionados con cuatro actividades como actividad en el baño, inquietud en la cama, actividad en la cocina y actividad en la sala de estar.

2.5 Reconocimiento de anomalía

El reconocimiento de actividades humanas también está siendo aplicado en el desarrollo de sistemas para reconocer anomalías, eventos adversos, comportamientos desviados a los habituales en una persona, para la provisión de intervenciones efectivas, oportunas y específicas de acuerdo con los objetivos de sus aplicaciones. Un desafío existente identificado en el área se debe a que existen comportamientos anormales que todavía no son cuantificables o son difíciles de observar.

En el artículo (Xiang & Gong, 2005), se utilizan características visuales ruidosas utilizadas para la representación del comportamiento. En esta investigación se define la anormalidad como patrones de comportamiento atípicos que no están representados por muestras suficientes en un conjunto de datos de entrenamiento. En la fase de experimentación se montó una cámara de CCTV en el techo de un pasillo de entrada a la oficina, monitoreando a las personas que entraban y salían del área, se recopiló un conjunto de datos durante cinco días diferentes que consta de seis horas de video, dados estos segmentos de video.

En el artículo (Vanus, 2017), se enfocan en monitorear el cambio en la rutina diaria de una persona que vive en una casa inteligente utilizando el análisis a largo plazo de las actividades realizadas. Primero se lleva a cabo el reconocimiento de actividad, se analiza la rutina diaria del ocupante explotando el conjunto de actividades de un día realizadas durante un período, con lo que el sistema es capaz de identificar una rutina normal, inusual y sospechosa.

En el artículo (Yin, Yang, & Pan, 2008) se define una actividad anormal como “eventos que ocurren raramente y no se han esperado de antemano”; su enfoque está basado en la observación de que no es realista obtener una gran cantidad de datos de entrenamiento para actividades anormales. Proponen dos fases para la detección de la actividad anormal. En la primera fase, se construye una máquina de vectores de soporte de una clase, basada únicamente en actividades normales, que puede filtrar las actividades que tienen una probabilidad muy alta de ser normales. Los rastros sospechosos se transmiten a una fase secundaria para una mayor detección.

En el artículo (Duong, et al., 2005), el sistema se diseñó para aprender, comprender y construir automáticamente un modelo de actividades de la vida diaria (ADL) mediante la observación de lo que suele hacer el ocupante estudiado durante la rutina del día, después el sistema utiliza su conocimiento aprendido para monitorear la actividad actual de la persona y detectar si existe alguna desviación de los patrones de actividad normales y alertar al cuidador si es necesario.

2.6 Sensores

Con el continuo avance de la tecnología, una de las fuentes de datos más utilizadas en los sistemas de reconocimiento y clasificación de actividades humanas, son los sensores. Los sensores portátiles se pueden conectar a diferentes partes del cuerpo y pueden medir alguna característica de interés o detectar algún tipo actividad. Los sensores mayormente utilizados para este fin, son los acelerómetros (Chen, et al., 2008), giroscopios (Mahmud, et al., 2020), los sensores de presión y los sensores de temperatura.

En el artículo (Lara, et al., 2012), se desarrolló un sistema basado en aceleración y señales fisiológicas para reconocer automáticamente actividades físicas como correr, sentarse, ascender y descender. Este sistema requiere un dispositivo sensor y un teléfono móvil, logrando 95.7% de precisión general. Los autores muestran que la posición del acelerómetro influye en la precisión del sistema. Las actividades descendiendo y caminando se reconocen sólo con acelerómetros, corriendo, sentado y subiendo escaleras con el uso de signos vitales.

En el artículo (Riboni & Bettini, 2011), un sistema desarrollado en Android es capaz de reconocer actividades de deambulación, así como cepillarse los dientes, pasear y escribir en la pizarra. Los datos se recopilan de dos sensores de acelerómetro; uno en la muñeca y otro en el teléfono celular, así como GPS del teléfono celular. Su evaluación determina que el sistema es capaz de lograr un 80%, 89% y 93% de precisión.

En el artículo (Maurer, et al., 2006), con el uso del sensor acelerómetro y sensor de luz, se realiza el reconocimiento y monitoreo de actividades como sentado, estar de pie, caminar y bajar escaleras usando múltiples sensores en diferentes posiciones del cuerpo. Se diseña una arquitectura con seis dispositivos *eWatch* para investigar la dependencia de la precisión de clasificación de *eWatch* en determinadas posiciones corporales

En el artículo (Mahmud, et al., 2020), se implementa una red neuronal de múltiples etapas a través de la integración de características multimodales de varios sensores portátiles. Los sensores utilizados son un giroscopio y 2 tipos de aceleradores, se logra un rendimiento promedio de 83.9% y precisión de 83.2%. Los autores sostienen que depender de un solo sensor limitará la diversidad de las características extraídas que se supone tienen problemas de reconocimiento en cuanto al tipo de actividad.

2.7 Antecedentes

En CENIDET, en el departamento de Ciencias Computacionales, existen trabajos de investigación que anteceden a esta tesis, es decir, aquellas investigaciones que logran la identificación, caracterización y seguimiento de una persona como objeto de estudio basado en Visión por Computadora para reconocer objetos en movimiento o acciones humanas.

Caracterización visual de movimientos sospechosos de personas en estacionamientos (Alcantar Juárez, 2013)

Se relaciona con este trabajo de tesis, ya que propone un sistema de vigilancia inteligente que es capaz de detectar y realizar el seguimiento de una o más de una persona a la vez, además de hacer la caracterización de cinco movimientos sospechosos: correr, levantar la mano o ambas manos, merodear, agacharse y acostarse.

Seguimiento y caracterización automática de la trayectoria de una persona en un entorno multi-cámaras de videovigilancia (Rosado Castellanos, 2019)

Otro trabajo relacionado, que parte de la identificación de una persona, se enfoca en clasificar la acción que realiza la persona semi-identificada por medio de características softbiométricas como: caminar derecho, caminar a la izquierda, girar a la derecha, quedarse parado.

Se enfoca al sistema de videovigilancia inteligente centrándose en la re-identificación de personas y análisis de trayectoria. El análisis de video, reconocimiento automático de las actividades humanas y el análisis de trayectoria es la fuente de información utilizada para identificar, de manera automática, comportamientos específicos o sospechosos.

Caracterización automática de una muestra de acciones en video (Dávila García, 2009)

En este trabajo de investigación se detecta, sigue, caracteriza y cuantifica de manera automática los cambios presentados por partes del cuerpo humano en movimiento al realizar las acciones de visualizar, aproximarse y tocar una zona específica, se tiene la capacidad de caracterizar numéricamente una muestra de acciones humanas en video de manera automática.

Se aborda el análisis del movimiento del cuerpo del ser humano ante una situación específica de interacción con objetos, en los movimientos de la cabeza, brazos y manos, así como la localización de la persona respecto al objeto.

2.8 Discusión

Al inicio del capítulo se abordan investigaciones que realizan el reconocimiento de actividades humanas, se mencionan las técnicas utilizadas y los objetivos de cada investigación mencionada.

En seguida se presentan investigaciones que realizan el reconocimiento de actividades, únicamente con técnicas de visión por computadora, donde se implementa la resta de fondo, características espacio temporales, hasta el uso de dispositivos más avanzados como Microsoft Kinect® para el seguimiento del esqueleto humano.

Posteriormente, se pueden encontrar investigaciones que han implementado la estimación de la pose humana para el reconocimiento de actividades deportivas, monitoreo de la vida diaria y reconocimiento de lenguaje de señas, algunas de las técnicas utilizadas son la implementación de *Open Pose* y *MediaPipe Pose*.

Avanzando en el capítulo, se presentan las investigaciones que realizan el monitoreo continuo de las actividades diarias que una persona realiza, para el modelado de comportamiento que puede ser implementado en aplicaciones para la vida asistida, medicina, monitoreo de adultos mayores, cuyo objetivo es obtener inferencias conductuales y predicciones sobre un comportamiento.

Finalmente, se presentan investigaciones que realizan el reconocimiento de patrones anómalos en el comportamiento de una persona. Además, se presentan algunas investigaciones que usan sensores portátiles según los objetivos de su investigación. Al final del capítulo se presentan antecedentes al tema de investigación de esta tesis.

CAPÍTULO 3

PROPUESTA DE SOLUCIÓN

El presente capítulo describe la propuesta de solución para alcanzar los objetivos de esta investigación, además se mencionan los alcances y limitaciones a los que se encuentra sujeto el desarrollo de este modelo para la detección de anomalías en el comportamiento de una persona. Se detalla la justificación y beneficios que se obtienen como resultado con esta investigación.

En este capítulo, además, se presenta un modelado conceptual de la solución con una breve descripción de los componentes que forman parte de este modelo. Al final, se enlistan las variables de control que son utilizadas.

3.1 Planteamiento de solución

El reconocimiento de anomalías en el comportamiento de personas es una investigación activa en el área de reconocimiento de actividades humanas (Duong, et al., 2005), (Alshurafa, et al, 2017).

En este trabajo de tesis se propone un modelo para el reconocimiento de anomalías en el comportamiento de una persona, por medio de un sistema conformado por cuatro cámaras ubicadas en la cocina, comedor, sala y dormitorio de un hogar y un teléfono móvil inteligente

3.1.1 Objetivo

Diseñar un sistema que integre cuatro cámaras fijas RGB y utilice los sensores disponibles en un dispositivo móvil para caracterizar, reconocer o identificar 16 actividades humanas. El modelo deberá monitorear la actividad del portador y realizar el reconocimiento de patrones anómalos en el comportamiento del humano estudiado.

Objetivos específicos

1. Desarrollar un sistema que procese los datos obtenidos de las cámaras fijas y el teléfono inteligente móvil, así como evaluar y seleccionar una metodología para el HAR.
2. Adquirir y monitorear datos del movimiento o actividades diarias de una persona.
3. Aprender con los datos monitoreados, a través de un modelo de *Machine Learning*, el comportamiento cotidiano de una persona.
4. Reconocer patrones anómalos en el comportamiento de una persona, con respecto al comportamiento aprendido.

3.1.2 Alcances y limitaciones

Alcances

- El modelo de visión por cámara puede personalizarse para identificar a cualquier persona dentro del rango de visión.
- El modelo de visión por cámara está diseñado para identificar a una sola persona de interés.
- El modelo de arquitectura de cámara de visión se adaptó para utilizar cuatro cámaras RGB.
- El sistema puede adquirir datos de movimiento y acciones humanas.
- Se espera que el sistema aprenda y reconozca 16 actividades humanas.
- El sistema puede recolectar datos de entrenamiento predefinidos para cada persona, con la viabilidad de implementar periodos semanales, mensuales, trimestrales, el periodo adecuado para la fase de entrenamiento será definido y seleccionado.
- El sistema se implementó en un fondo simple tolerando ligeras variaciones en el fondo.
- El sistema es capaz de reconocer cuando existe una variación en el comportamiento de una persona.

Limitaciones

- No es compromiso que el sistema recopile datos de más de una persona a la vez.
- Este estudio no es capaz de evaluar periodos de recolección de datos para el entrenamiento del sistema que comprenda una duración mayor que represente un riesgo para alcanzar los objetivos de esta investigación.
- La recopilación de los datos obtenidos de los sensores se realizó dentro del mismo rango de visión de las cámaras del sistema.

3.1.3 Justificación y beneficios

Dada la importancia del reconocimiento de actividad humana en las diferentes áreas de investigación y desarrollo, se puede pensar que hay un creciente interés en desarrollar aplicaciones que permitan el monitoreo y la detección de anomalías en el comportamiento de una persona (Skubic, Guevara, et al., 2015).

En la literatura se han desarrollado múltiples metodologías donde se realiza el monitoreo de actividades de la vida diaria para reconocer patrones de comportamiento (Pavel, et al., 2015). El desarrollo de estos sistemas ha permitido aplicarse en el área de medicina, vida asistida, monitoreo de adultos mayores (Thummala & Pumrin, 2020), entre otros. Sin embargo, se considera que aún existen desafíos en el área de reconocimiento de patrones de comportamiento (Alshurafa, et al., 2017).

Aunque se han diseñado metodologías eficientes para el reconocimiento de anomalías, no se encontró ningún método en la literatura revisada, que realice el reconocimiento de actividades humanas con la estimación de la pose (Cao, Hidalgo, et al., 2021) la interacción con objetos del entorno y se enfoque en el reconocimiento de patrones anómalos en el comportamiento de una persona.

En conclusión, se puede considerar como una gran oportunidad para innovar en el campo de reconocimiento de patrones de comportamiento humano, que podría llevar al desarrollo de modelos enfocados al reconocimiento de anomalías en el comportamiento de una persona.

3.2 Modelo conceptual de la solución

En esta sección del documento se presenta en la Figura 3.1, el diagrama de componentes que representa el funcionamiento de la solución propuesta para la investigación. Además, se detalla una descripción de los elementos del modelo conceptual.

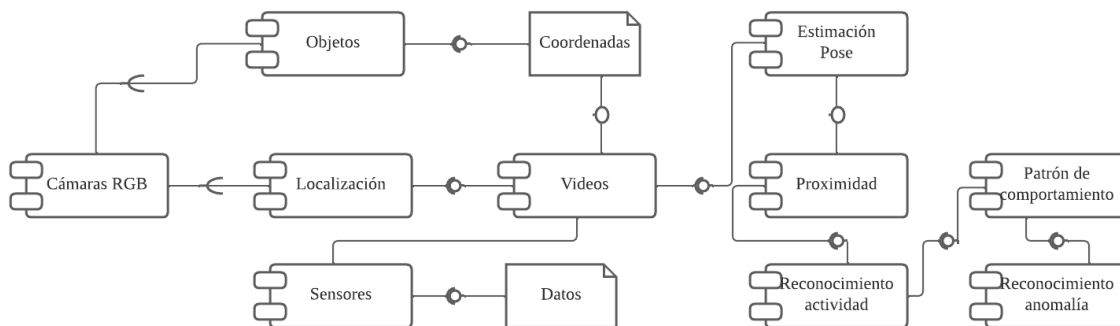


Figura 3.1. Diagrama de componentes

3.3 Descripción de los elementos del modelo conceptual

En esta sección se detallan los elementos que conforman al modelo propuesto y la funcionalidad que cumplen en el modelo.

Datos de cámaras RGB

Cuatro cámaras RGB son distribuidas en cuatro habitaciones de un hogar, cocina, comedor, sala de estar y dormitorio. Las cámaras utilizadas son de la marca Dahua, modelo DH – IPC – HFW1431SN – 02808 – S4, resolución HD 720 (1280×720), 30 FPS.

Detección de objetos

Para la detección automática de objetos, se está usando *YOLO V3* (Redmon, et al., 2018), un modelo previamente entrenado, con una lista de 80 clases de objetos que el modelo puede detectar. De los objetos detectados se realiza una selección de aquellos con los que la persona interactúa dentro de su rutina.

Localización de una persona

Para realizar la localización de una persona dentro una habitación en toda la distribución del hogar se probaron los siguientes métodos de localización:

1. Sensor RFID.
2. Pulsadores eléctricos.
3. Aplicación Móvil.

Grabación de video

El sistema realiza la grabación de video con una duración aproximada a tres minutos de la secuencia de actividades que una persona está realizando, dentro de la habitación en la que fue localizada.

Datos sensores

Se desarrolló una aplicación móvil en *Android Studio*, que permite adquirir datos de los sensores acelerómetro y giroscopio a través de un dispositivo móvil. La recopilación de los datos de los sensores es utilizada para monitorear el desplazamiento de una persona en los puntos ciegos del escenario, es decir, cuando la persona se desplaza de una habitación a otra sin poder ser visualizado por las cámaras.

Estimación de la pose humana

La estimación de la pose humana (*Human Pose Estimation*) es el proceso de predecir las posiciones 2D o 3D de cada una de las partes del cuerpo de una o varias personas que se

encuentren presentes en imágenes o video. Los métodos evaluados para la estimación de la pose humana son los siguientes:

- *Open Pose*
- *MediaPipe Pose*

Medida de proximidad

Al procesar un video se realiza el cálculo de una medida de proximidad, que indica una medida de distancia con respecto a los objetos previamente detectados en cada una de las habitaciones y la posición de las articulaciones, resultado de la estimación de la pose humana para el reconocimiento de las actividades que hace una persona.

Se calcula la distancia *Manhattan*, utilizando las coordenadas cartesianas (x,y) de las articulaciones, y el centro (x, y) del objeto detectado por YOLO V3, para la identificación de las actividades realizadas por una persona.

Reconocimiento de actividades

La estimación de la postura de una persona y el cálculo de proximidad dan lugar al reconocimiento de actividades, se reconoce un total de 16 actividades que una persona puede desempeñar en el hogar.

Patrón de comportamiento

El modelo procesa los videos resultados de una rutina diaria de actividades y a partir de la información obtenida es capaz de generar grafos dirigidos para cada una de las habitaciones. Los nodos o estados representan las actividades y las aristas trazan las transiciones de una actividad a otra. El recorrido de un grafo estaría modelando la información de la secuencia de actividades que se realizan durante una rutina diaria.

Al agrupar los recorridos de los 4 grafos que son generados por el modelo, se obtiene como resultado un grafo general de todas las habitaciones que representa los patrones de comportamiento que se infieren de procesar una rutina diaria.

Reconocimiento de la anomalía

El modelo propuesto utiliza máquinas de estado finito (*Finite State Machine*), también llamado autómata finito, como reconocedoras, a través de un alfabeto, transiciones, un estado inicial y conjunto de estados finales, se puede obtener un patrón resultado de comenzar en un estado inicial y finalizar con un estado de aceptación o no aceptación.

El patrón contiene información de la secuencia de actividades realizadas por una persona durante su rutina, que permite modelar su comportamiento. La máquina de estado finito define los estados de aceptación a través del entrenamiento del modelo y permite reconocer si el patrón introducido es aceptado “normal” o rechazado “anormal”, además de reconocer si existen anomalías, al día, al periodo, o si se trata de un patrón inhabitual para una persona.

3.4 Variables / criterios de control

En esta sección del documento se hace mención de la definición de las variables y criterios de control que se establecieron para el correcto funcionamiento del modelo.

- **Umbral de probabilidad para el objeto predicho:** Es considerado en la detección de objetos para hacer la selección de los objetos que presentaron los mejores resultados al ser detectados por *Yolo v3*.
- **Frecuencia de muestreo en sensores:** Se decidió adquirir los datos proporcionados por el acelerómetro y giroscopio a una frecuencia de 60 Hz.
- **Parámetro de visibilidad de la estimación de la pose humana:** El método *MediaPipe pose* predice la posición de 33 puntos de referencia del cuerpo humano y proporciona un parámetro de visibilidad que representa la confianza de detección de cada articulación de manera individual. Este parámetro es utilizado por el modelo propuesto para realizar el reconocimiento de una actividad.
- **Transiciones válidas en máquina de estado finito:** Condición que se debe de cumplir al introducir un patrón de comportamiento a la máquina de estado finito.
- **Estados de aceptación en máquina de estados finitos:** Condición de una máquina de estado finito que va a permitir reconocer si un patrón alcanza un estado de aceptación, o no aceptación, y por consecuencia se realice el reconocimiento de una anomalía.

CAPÍTULO 4

Análisis, diseño e implementación del sistema

En este capítulo, se describe el análisis y diseño del proceso para realizar la implementación del modelo para la detección de anomalías en el comportamiento de una persona. El proceso incluye las siguientes etapas: localización de una persona, estimación de la pose humana, reconocimiento de objetos, reconocimiento de actividad, recolección de datos de sensores, modelado de una rutina, finalmente el reconocimiento de anomalía.

4.1 Arquitectura del sistema

En la Figura 4.1 se muestra, de manera general, las fases que incluye el modelo para la detección de anomalías.

Fase de adquisición: se refiere a obtener imágenes de las habitaciones y grabaciones de las actividades diarias de la rutina de una persona. En esta misma fase, se realiza la detección de objetos de manera automática, las coordenadas obtenidas de los objetos se le proporcionan al modelo para etiquetar las zonas donde se reconoció un objeto, además se requiere realizar una localización de la persona, para conocer en cuál de las habitaciones se encuentra.

Fase de localización: se experimentó con sensores adicionales y un teléfono inteligente para localizar a una persona dentro de una habitación, si la localización se ha logrado, el modelo identifica cuál es la habitación que debe empezar a grabar. Se realiza la grabación de video. Asimismo, la persona localizada porta un teléfono inteligente a la altura de su cintura en el bolsillo de su pantalón, que permite monitorear los datos del sensor, acelerómetro y giroscopio.

Fase de reconocimiento de actividad: Se procesa un video para obtener la estimación de la pose humana, usando la librería *MediaPipe Pose*, seguido de calcular una medida que indica una proximidad a ciertos objetos detectados automáticamente. Este cálculo se realiza con las coordenadas cartesianas (x,y) del centro de un objeto previamente detectado y la posición de las articulaciones del cuerpo de una persona. Con estos parámetros se calcula la distancia

Manhattan y se reconoce la actividad asociada entre el objeto y la cercanía de una persona. En total se reconocen 16 actividades de la vida diaria.

Fase de reconocimiento de anomalía: el resultado de procesar un video es el reconocimiento de la secuencia de actividades que realiza una persona durante la grabación de este video. El modelo genera la unión de las aristas de un grafo, según la secuencia de actividades reconocida.

Por lo tanto, se dice que el modelo crea grafos dirigidos para cada una de las habitaciones, además de un grafo general que contiene la información de todas las actividades posibles (16 actividades), y que representa la secuencia de actividades que una persona ha realizado durante su rutina. Cuando se han formado los grafos dirigidos a partir de lo reconocido en los videos procesados, se han generado patrones de comportamiento de la persona que fue monitoreada. El siguiente proceso es utilizar máquinas de estado finito como reconocedoras de cadenas, a las que se provee de la información de los grafos dirigidos construidos. Las máquinas de estado finito como reconocedoras permitirán clasificar un patrón “normal” o “anormal”, según la cadena de caracteres que se introduzca.

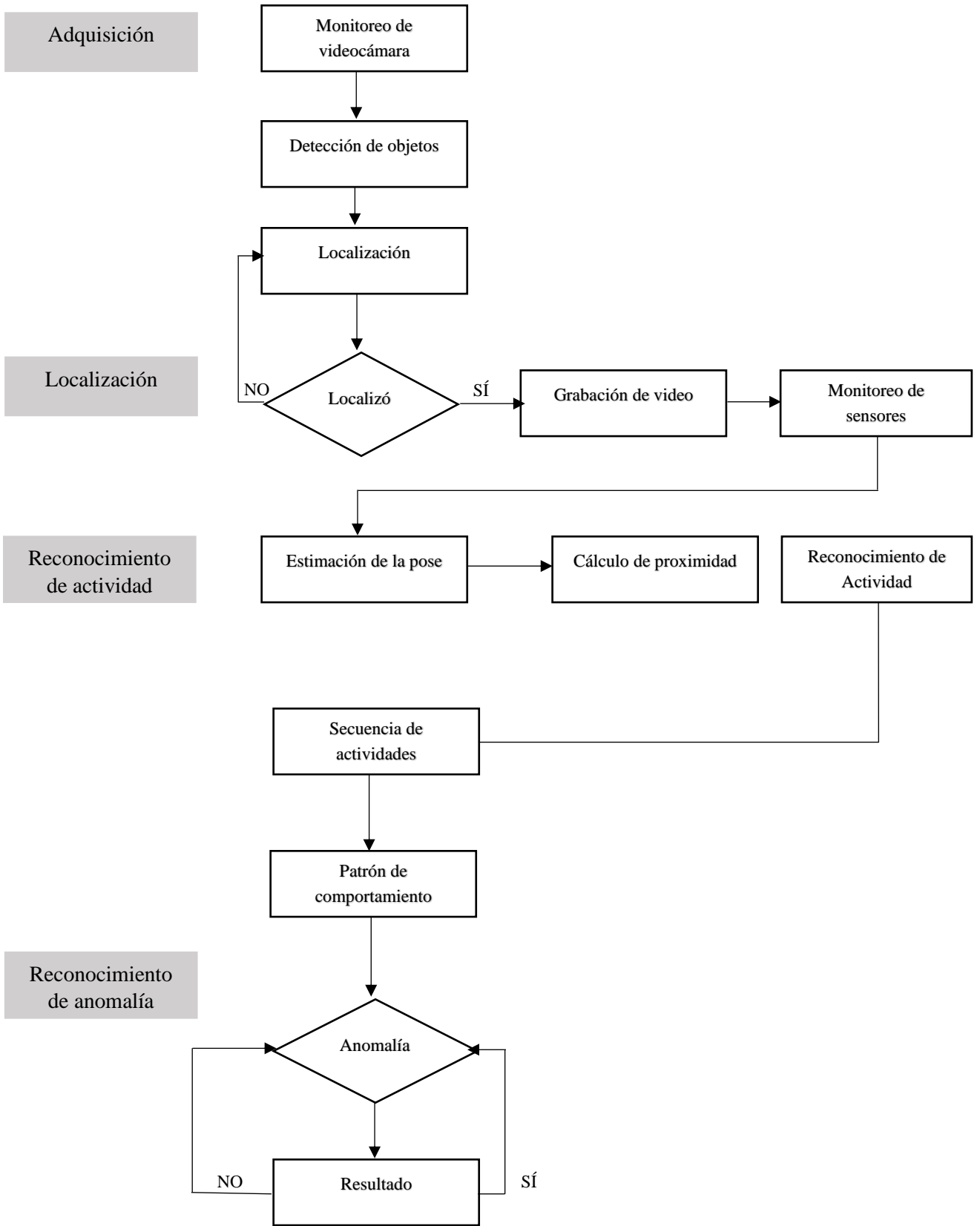


Figura 4.1. Arquitectura del sistema

4.2 Etapa de localización de una persona

La etapa de localización de una persona se refiere a conocer la ubicación de la persona dentro del espacio de distribución.

La arquitectura del sistema utiliza cuatro cámaras RGB, en una distribución de cuatro habitaciones (cocina, comedor, sala de estar y dormitorio) dentro de un escenario seleccionado como hogar, consultar Figura 4.2 para ubicar la distribución de cámaras.

Una persona puede entrar y salir de una habitación en todo el espacio de distribución, se seleccionaron solamente cuatro habitaciones importantes para realizar la rutina diaria de una persona.

En el estado del arte, algunas de las técnicas para la localización de una persona en la escena, fue la resta de imagen de fondo 2D y la estimación de un cuadro delimitador para rastrear el cuerpo en un sistema de vigilancia (Babiker, et al., 2017). Otro método en un sistema de monitoreo de adultos mayores fue la cuantificación del movimiento de una persona y para analizar la forma humana se implementó la resta de fondo (Thummala & Pumrin, 2020).

En esta investigación, para poder localizar la ubicación de una persona se probaron los siguientes métodos de localización: sensor RFID, pulsadores eléctricos y uso de aplicación móvil, la cual permitirá localizar a la persona de interés dentro de las cuatro habitaciones seleccionadas. Si el sistema ha localizado a una persona, comenzará la grabación de un video de la habitación en la que la persona fue localizada.

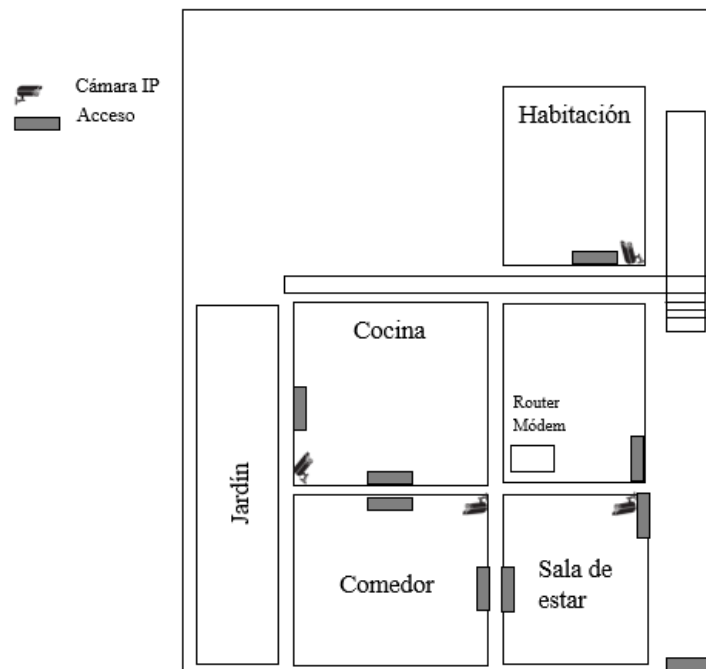


Figura 4.2. Diagrama de distribución de cámaras RGB

4.2.1 Sensor RFID

La tecnología RFID (identificación por radiofrecuencia) permite realizar funciones de trazabilidad, identificación de objetos (seguimiento) y control de acceso. La información se almacena en una memoria a la que se puede acceder mediante un simple enlace de radiofrecuencia. Esta memoria es en forma de una etiqueta electrónica, que contiene una antena y un circuito integrado (Barrett, S. F, 2013). La etiqueta contiene la información asociada al objeto al que está fijada. Cuando una etiqueta entra en el campo generado por el lector, detecta la señal e intercambia los datos entre su memoria y el lector.

La implementación de este sensor fue programada, con el controlador Arduino nano y las lecturas del sensor se transfieren mediante el uso de un módulo wifi esp8266. De esta forma es posible que una persona porte un lector RFID (ver Figura 4.3), y se coloquen *tags* (etiquetas en forma de llavero) o tarjetas en las cuatro habitaciones del hogar, (ver Figura 4.4), que contienen antenas para permitirles recibir y responder a peticiones por radiofrecuencia desde un emisor – receptor RFID, (ver Figura 4.5).

Cada una de las etiquetas está asociada a un identificador de habitación, ocasionando que la lectura del identificador pueda localizar a la persona dentro una habitación específica.

Un ejemplo para la localización, es que una persona porta el lector RFID y se aproxima a los tags (etiqueta) colocada en el acceso de la habitación “cocina”. Como resultado, el lector RFID envía la lectura obtenida y el sistema localiza a la persona dentro de la habitación “cocina”.

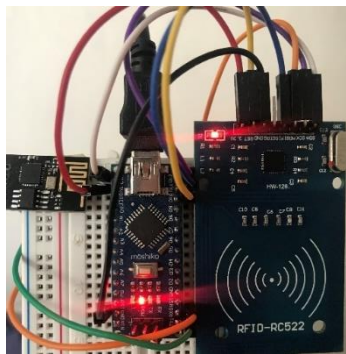


Figura 4.3. Lector RFID

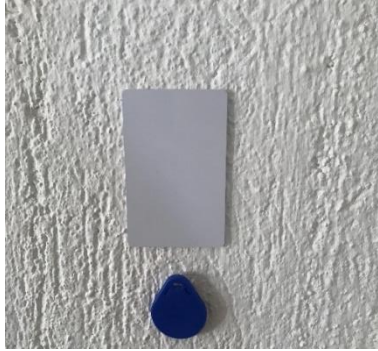


Figura 4.4. Tarjetas (tags)

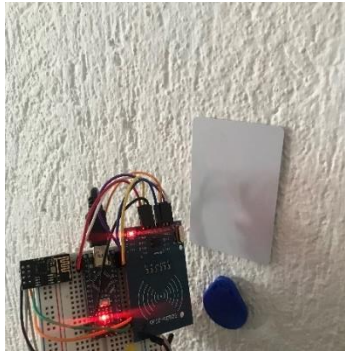


Figura 4.5. Lector RFID y tarjetas

4.2.2 Pulsadores eléctricos

Un pulsador eléctrico o botón pulsador es un componente que permite o impide el paso de la corriente eléctrica cuando se aprieta o se pulsa.

Se implementó el uso de pulsadores eléctricos como el que se muestra en la Figura 4.6, junto con un módulo wifi esp8266, utilizando el controlador Arduino uno, programando un pulsador por cada una de las habitaciones del hogar. El pulsador se abre o cierra cuando la persona monitoreada lo presiona o mantiene presionado y se logra la localización de la persona en la habitación debido a que cada pulsador está asociado a una habitación del hogar.

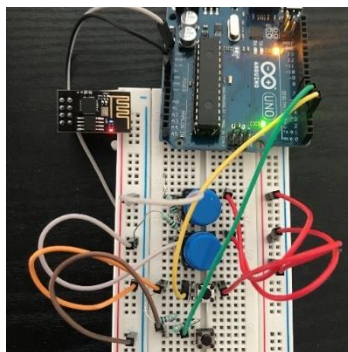


Figura 4.6. Pulsadores eléctricos

Un ejemplo para la localización, es que una persona porta el control de pulsadores ver Figura 4.7 y presiona el primer pulsador que este asociado a la habitación “cocina”, el segundo al “comedor”, el tercero a la “sala de estar” y por último el botón asociado a la habitación “dormitorio”. Cuando se ha presionado el botón, el sistema es capaz de localizar a la persona según el botón que se haya presionado.

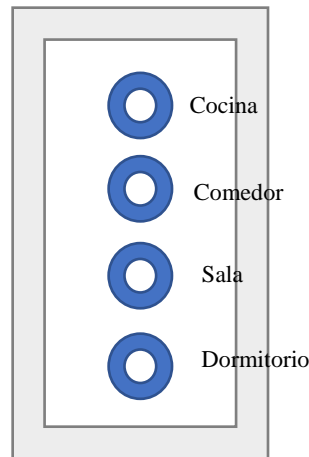


Figura 4.7. Diseño de control de pulsadores eléctricos

4.2.3 App móvil

Se desarrolló una aplicación móvil intuitiva ver Figura 4.8, que permite a una persona seleccionar un botón asociado a la habitación en la que se encuentra, estos botones están etiquetados con el nombre de la habitación que le corresponde. Esta aplicación está desarrollada en Android Studio (Novac, et al., 2017), programada con lenguaje Java, disponible para teléfonos y *tablets* inteligentes.

Un ejemplo para la localización es portar un teléfono o *tablet* inteligente, abrir la aplicación de localización, una persona deberá seleccionar en la aplicación el botón que tenga el nombre de la habitación en la que se encuentra. Si una persona se encuentra en la habitación “dormitorio” deberá seleccionar este botón y por consecuencia el sistema podrá localizar a la persona dentro de la habitación “dormitorio”.



Figura 4.8. Interfaz gráfica de aplicación móvil

4.2.4 Adquisición de video

La adquisición de video graba la actividad o actividades que realiza una persona, después de haber sido localizada por el sistema.

Las cuatro cámaras se encuentran encendidas y monitoreando las cuatro habitaciones en las que fueron ubicadas dentro del hogar. El sistema se encuentra a la espera de localizar a una persona. Al implementar un método de localización de una persona, se obtiene la información de la habitación que se tiene que monitorear y grabar.

En la habitación donde se ha dado la localización de una persona, el sistema comienza la grabación de las actividades que la persona está realizando. La grabación es almacenada para posteriormente utilizar los videos para su procesamiento.

4.3 Etapa de estimación de la pose humana

La etapa de la estimación de la pose humana consiste en procesar mediante el uso de una librería un video que contiene una actividad o secuencia de actividades realizadas por una persona que está siendo monitoreada por el sistema, con el objetivo de obtener la estimación de la pose.

La estimación de la pose humana, es el proceso de obtener la configuración espacial de las partes del cuerpo en las imágenes o videos proporcionados por la cámara RGB.

Las siguientes investigaciones forman parte del estado del arte, que implementaron la estimación de la pose humana haciendo uso de diferentes métodos. En un sistema que registra actividades de gimnasio se implementó Open Pose con el objetivo de contabilizar a través de características basadas en los puntos del esqueleto (Ganesh, Idgahi, et al., 2020). Otro sistema propone asistencia para adultos mayores para la detección de caídas, el sistema realiza el seguimiento del esqueleto proporcionado por Microsoft Kinect® (Mettel, et al., 2019).

En esta sección se presentan los métodos de estimación de la pose humana implementados en esta investigación.

4.3.1 Open Pose

Open Pose (Cao, Hidalgo, et al., 2021), presenta un enfoque en tiempo real para detectar la pose 2D de varias personas en una imagen. El método utiliza una representación no paramétrica, a la que se refiere como *Part Affinity Fields* (PAF), para aprender a asociar partes del cuerpo con individuos en la imagen en tiempo real, independientemente del número de personas en la imagen. *Open Pose*, es el primer sistema de código abierto en tiempo real para la detección de poses 2D de varias personas, incluidos los puntos clave del cuerpo, el pie, la mano y la cara.

La Figura 4.9 ilustra el proceso general del método propuesto en *Open Pose*. El sistema toma como entrada, una imagen en color de tamaño $w \times h$ y produce las ubicaciones 2D de los puntos clave anatómicos para cada persona ver Figura 4.9 (e).

Primero, una red de retroalimentación predice un conjunto de mapas de confianza 2D de las ubicaciones de las partes del cuerpo consultar Figura 4.9 (b) y un conjunto de campos vectoriales 2D, L de campos de afinidad de partes (PAF), que codifican el grado de asociación entre partes ver Figura 4.9 (c).

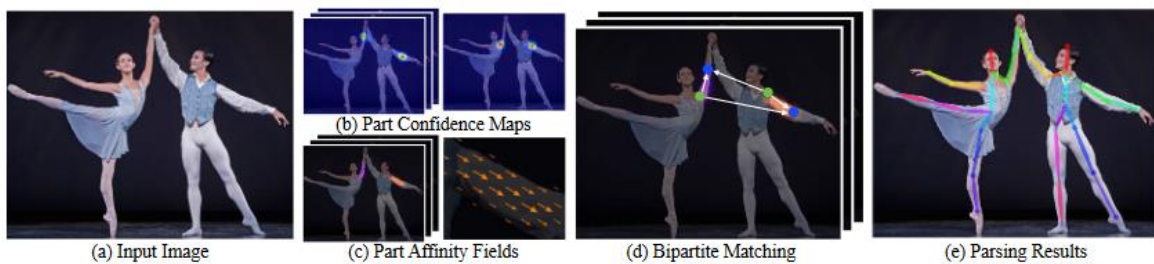


Figura 4.9. Proceso del método propuesto en *Open Pose* (Cao, Hidalgo, et al., 2021)

El conjunto $S = (S_1, S_2, \dots, S_J)$ tiene J mapas de confianza, uno por parte, donde $S_j \in R^{w \times h}$, $j \in \{1 \dots J\}$. El conjunto $L = (L_1, L_2, \dots, L_C)$ tiene C campos de un vector, uno por rama, donde $L_C \in R^{w \times h \times 2}$, $C \in \{1 \dots C\}$. Se refiere a los pares de partes como miembros para mayor claridad, pero algunos pares no son miembros humanos, por ejemplo, la cara.

Cada ubicación de la imagen en L_C codifica un vector 2D. Finalmente, los mapas de confianza y los PAF se analizan mediante inferencias codiciosas para generar los puntos clave 2D para todas las personas ver Figura 4.10.



Figura 4.10. Codificar un vector 2D Open Pose (Cao, Hidalgo, et al., 2021)

Para obtener la estimación de la pose humana, el sistema deberá procesar un video que contiene actividades diarias realizadas por una persona. El resultado de la estimación deberá ser como se muestra en la Figura 4.11.

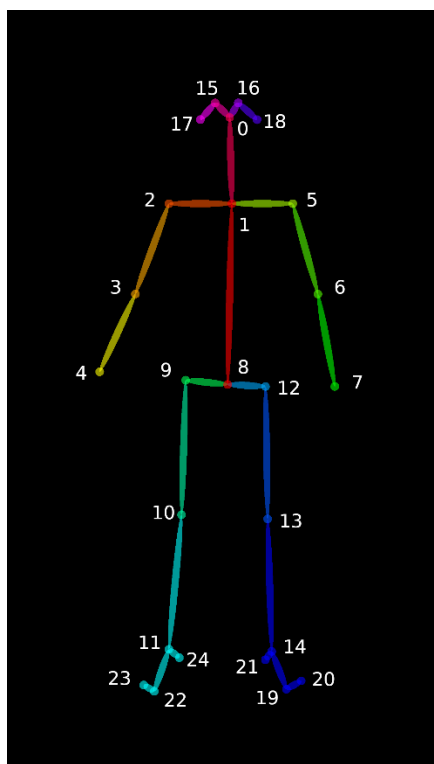


Figura 4.11. Resultado de estimación de la pose Open Pose (Cao, Hidalgo, et al., 2021)

4.3.2 MediaPipe Pose

MediaPipe Pose (Bazarevsky, et al., 2020) es una librería de aprendizaje automático de *Google*, para el seguimiento de posturas corporales de alta fidelidad, que infiere 33 puntos de referencia 3D y una máscara de segmentación de fondo en todo el cuerpo a partir de cuadros de video RGB utilizando la investigación *BlazePose*.

Esta librería ha sido implementada como en las investigaciones (Chundur., et al, 2021) y (Sreenath., et al, 2021). Los enfoques de vanguardia actuales se basan principalmente en entornos de escritorio potentes para la inferencia, mientras que este método logra un rendimiento en tiempo real en la mayoría de los teléfonos móviles, computadoras de escritorio / portátiles, en el lenguaje Python e incluso en la web.

Modelo de detección de persona / pose (Detector *BlazePose*)

El detector está inspirado en su propio modelo ligero llamado *BlazeFace*, utilizado en *MediaPipe Face Detection*, como un proxy para un detector de personas. Predice explícitamente dos puntos clave virtuales adicionales que describen firmemente el centro del cuerpo humano, la rotación y la escala como un círculo. Inspirándose en el hombre de Vitruvio de Leonardo, se predice el punto medio de las caderas de una persona, el radio de un círculo que circunscribe a toda la persona y el ángulo de inclinación de la línea que conecta los puntos medios del hombro y la cadera ver Figura 4.12. El modelo en *MediaPipe Pose* predice la ubicación de 33 puntos de pose, ver Figura 4.13.

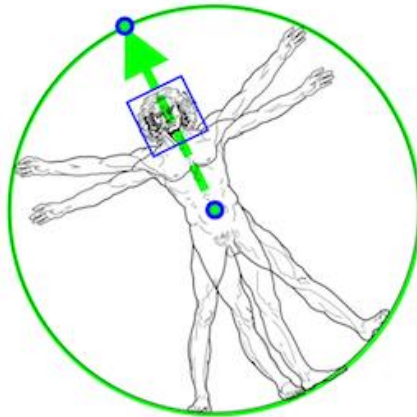


Figura 4.12. Hombre de Vitruvio alineado a través de dos puntos clave virtuales predichos por el detector *BlazePose* (Bazarevsky, et al., 2020).

Para la implementación de la estimación de la pose con la librería *MediaPipe Pose*, se programó un sistema en lenguaje Python, se proporcionan videos con duración aproximada a los tres minutos. El video incluye de actividades de la vida diaria realizadas por una persona. Los resultados de la estimación de la pose humana al procesar un video, tiene como resultado el cálculo de 33 puntos de referencia, consultar Figura 4.13.

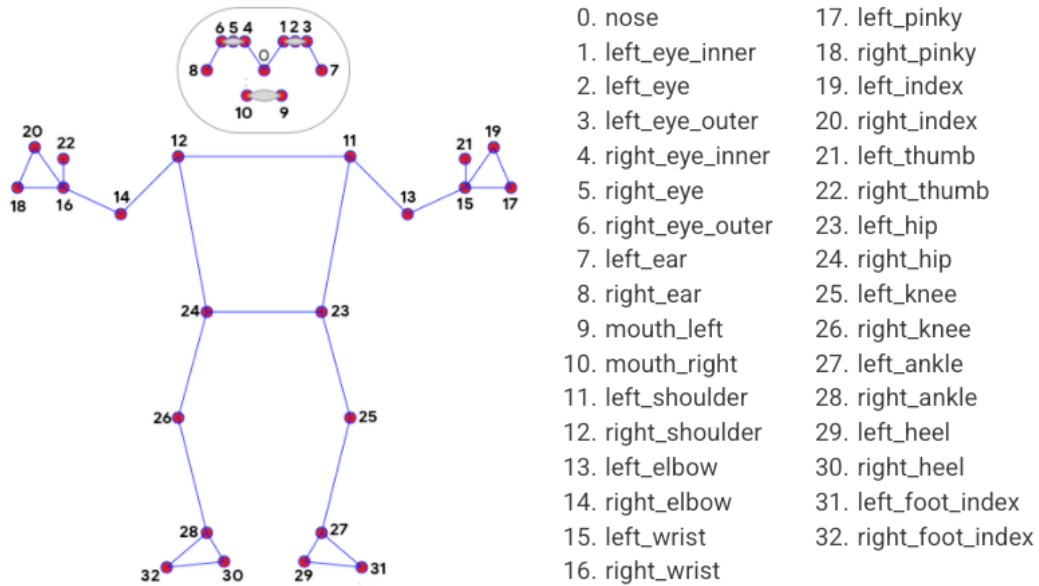


Figura 4.13. Puntos de referencia de la pose de MediaPipe Pose (Bazarevsky, et al., 2020).

4.4 Etapa de reconocimiento de objetos

El escenario utilizado para la investigación es un hogar, del cual se seleccionaron cuatro habitaciones para monitorear una rutina. En las habitaciones se encuentran objetos que son utilizados para el diario vivir.

En esta investigación se utiliza información de la ubicación de algunos objetos de interés, con los que una persona suele interactuar, al realizar una rutina diaria. Para esto es necesario realizar una etapa de detección de objetos de manera automática y obtener sus coordenadas.

Se estudió en la literatura una alternativa para realizar la detección de objetos, se encontró diferentes modelos como RetinaNet (Lin, et al., 2017), Google Spinet (Ganesh, et al., 2020), Detection Transformer (DETR) (Carion, et al., 2020) una solución propuesta por Facebook.

Con el objetivo de seleccionar un modelo que permita procesar una imagen y devuelva como resultado las coordenadas de un objeto previamente detectado. La solución óptima para resolver este proceso, es con la implementación de la red *YOLO V3* (Redmon, et al., 2018). Esta red hace una única entrada a la red convolucional y detecta todos los objetos para los que fue entrenada para clasificar.

Al ser un solo cálculo y sin necesidad de iterar, logra velocidades nunca antes alcanzadas con ordenadores que no tienen que ser tan potentes. Esto permite detección sobre video en tiempo real de cientos de objetos en simultáneo y hasta su ejecución en dispositivos móviles.

4.4.1 Yolov3

YOLO por sus siglas en inglés *You only look once* (Redmon, et al., 2018), es un modelo de detección de objetos en tiempo real de última generación, que puede ser aplicado a imágenes, videos o cámara web. Los sistemas de detección en versiones anteriores desarrollados por YOLO reutilizan los clasificadores o localizadores para realizar la detección.

Aplican el modelo a una imagen en múltiples ubicaciones y escalas. Las regiones de alta puntuación de la imagen se consideran detecciones. YOLOV3 (Redmon, et al., 2018), utiliza un enfoque totalmente diferente. Aplica a una única red neuronal a la imagen completa. Esta red divide la imagen en regiones y predice cuadros delimitadores y probabilidades para cada región. Estos cuadros delimitadores están ponderados por las probabilidades predichas.

YOLOV3 desarrolló una nueva red para realizar la extracción de características. La nueva red es un enfoque híbrido entre la red utilizado en YOLOv2, Darknet-19, la red que utiliza capas convolucionales sucesivas 3×3 y 1×1 pero ahora tiene algunas conexiones de atajos y es significativamente más grande. Tiene un total de 53 capas convolucionales consultar Figura 4.14.

La red es llamada entonces Darknet-53, esta nueva red es mucho más poderosa que Darknet19, pero aún más eficiente que ResNet-101 o ResNet-152. En el pasado, YOLO luchaba con objetos pequeños. Sin embargo, ahora se enfoca una reversión en esa tendencia. Con el nuevo predictor multiescala, YOLOv3 tiene relativamente alto rendimiento. Sin embargo, tiene comparativamente peor rendimiento en objetos de tamaño mediano y grande.

	Type	Filters	Size	Output
	Convolutional	32	3×3	256×256
	Convolutional	64	$3 \times 3 / 2$	128×128
1x	Convolutional	32	1×1	128×128
	Convolutional	64	3×3	
	Residual			
	Convolutional	128	$3 \times 3 / 2$	64×64
2x	Convolutional	64	1×1	64×64
	Convolutional	128	3×3	
	Residual			
	Convolutional	256	$3 \times 3 / 2$	32×32
8x	Convolutional	128	1×1	32×32
	Convolutional	256	3×3	
	Residual			
	Convolutional	512	$3 \times 3 / 2$	16×16
8x	Convolutional	256	1×1	16×16
	Convolutional	512	3×3	
	Residual			
	Convolutional	1024	$3 \times 3 / 2$	8×8
4x	Convolutional	512	1×1	8×8
	Convolutional	1024	3×3	
	Residual			
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figura 4.14. Estructura de la red Darknet-53 (Redmon, et al., 2018)

Se deberá ejecutar la librería con las imágenes de cada una de las habitaciones y obtener resultados acerca de los objetos detectados en cada habitación. Se proporciona una imagen

con dimensión de 1280×720 píxeles, de cada una de las cuatro habitaciones (cocina, comedor, sala y dormitorio) seleccionadas del hogar consultar Figura 4.15 - Figura 4.18. Se toman las coordenadas (x,y) de los objetos seleccionados y la configuración espacial, se utilizará en las siguientes etapas para procesar los videos del monitoreo de la rutina de una persona. Se configuró las coordenadas de los objetos detectados en el modelo que procesará los videos de monitoreo de la rutina de una persona.

Se pretende que el ambiente sea controlado, es decir, que los objetos permanezcan siempre en la misma ubicación.

En las siguientes imágenes se puede consultar las cuatro habitaciones del hogar, primero se observa la cocina Figura 4.15, segundo el comedor Figura 4.16, después la sala de estar Figura 4.17 y finalmente en la planta alta el dormitorio Figura 4.18.



Figura 4.15. Habitación Cocina



Figura 4.16. Habitación Dormitorio



Figura 4.17. Habitación Sala



Figura 4.18. Habitación Dormitorio

4.5 Etapa de recolección de datos de sensores

Para la implementación de los sensores portátiles hay factores que se deben considerar para obtener mejores resultados, desde la selección del tipo de sensor, tipo de actividades, ubicación del sensor, frecuencia de muestreo y aspectos técnicos. En el reconocimiento y monitoreo de actividades (Maurer, et al., 2006), sus resultados comprueban una precisión de reconocimiento mayor al usar dispositivos *eWatch* en la muñeca 87% y en el cinturón con 84.2%.

En otra investigación se destaca que el uso de los sensores acelerómetro y giroscopio, son sensores indicados para monitorear actividades del tipo ambulatorio (Mahmud, et al., 2020), además de proponer el uso de características multimodales con el uso de varios sensores portátiles que resulta en un desempeño mayor en todas las métricas de evaluación, además de que depender de un solo sensor limitará la diversidad de las características extraídas.

En la etapa de recolección de datos de sensores se desarrolló una aplicación móvil en *Android Studio* (consultar interfaz gráfica de la aplicación en la Figura 4.19), que permite obtener datos de los ejes de los sensores acelerómetro y giroscopio a través de un dispositivo móvil teléfono inteligente portado en el bolsillo del pantalón de la persona monitoreada.

Dispositivo implementado para esta investigación:

- Teléfono inteligente Moto G8.



Figura 4.19. Interfaz gráfica de aplicación móvil

Para la implementación, los datos de los sensores son capturados a una frecuencia de muestreo de 60 Hz, para el preprocesamiento de los datos de los sensores, se realizó una investigación en la literatura para poder seleccionar que técnica de preprocesamiento o filtrado de los datos implementar en esta investigación, ver Tabla 4.1.

Tabla 4.1. Preprocesamiento de los datos del sensor

<i>Dataset</i>	Frecuencia	Filtro	Características
ADL (Bruno, et al., 2013)	32Hz	Filtro mediano	Modelo GMM
UCI HAR (García Gonzales, et al., 2020)	50Hz	Filtro mediano	Variables del dominio del tiempo y frecuencia

Tabla 4.1. Preprocesamiento de los datos del sensor (Continuación)

<i>Dataset</i>	Frecuencia	Filtro	Características
UTDMHAD (C. Chen, et al., 2015)	30Hz	Filtro pasa bajo	Características estadísticas
WISDM (Weiss, et al., 2008)	20Hz	Filtro Butterworth de paso bajo	Características estadísticas

Filtro Pasa Bajo

El filtro pasa bajo corresponde a un filtro cuya característica principal es el de atenuar las frecuencias más altas y dejar pasar solo las frecuencias más bajas. Es un filtro óptimo que, al aplicarlo a una lectura de algún sensor, tiende a contener ruido en sus valores o cuyos valores obtenidos no son completamente estables, como, por ejemplo, un sensor de vibración, un sensor de aceleración o sensor de giroscopio utilizados en la investigación. Se implementa un filtro pasa bajo a cada una de las señales de:

- Sensor acelerómetro: mide m/s^2 la fuerza de aceleración que se aplica a un dispositivo en los tres ejes físicos (x, y, z).
- Sensor giroscopio: mide en rad/s la velocidad de rotación de un dispositivo alrededor de cada uno de los tres ejes físicos (x, y, z).

La recopilación de los datos de los sensores es utilizada para monitorear el desplazamiento de una persona en los puntos ciegos del escenario, es decir, cuando la persona se desplaza de una habitación a otra sin poder ser visualizado por las cámaras.

4.6 Etapa de Reconocimiento de actividad

El modelo de detección de anomalías ejecuta el reconocimiento de las actividades diarias realizadas por una persona, implementando la siguiente metodología:

1. Adquirir videos de cuatro cámaras RGB.
2. Realizar la detección de objetos en video.
3. Realizar la localización de una persona y grabar video.
4. Adquirir datos del sensor acelerómetro y sensor giroscopio.
5. Realizar la estimación de la pose humana.
6. Realizar el cálculo de medida de proximidad.

Para la etapa de reconocimiento de actividad en un video, es necesario ejecutar las fases de la metodología mencionada. Se deberá adquirir la grabación de un video de cualquiera de las habitaciones seleccionadas, previamente realizar la detección de objetos utilizando una librería y configurando las coordenadas donde se detectaron los objetos de interés.

La siguiente etapa consiste en adquirir los datos de los ejes de los sensores acelerómetro y giroscopio del desplazamiento de la persona al realizar las actividades de su rutina diaria.

En seguida, se realizó el procesamiento del video para obtener la estimación de la pose humana, al mismo tiempo que se está estimando la pose humana, el modelo deberá ser capaz de calcular una cercanía con respecto a los objetos detectados en la habitación.

La medida de proximidad permitirá asociar entre la distancia de una persona y un objeto una etiqueta de actividad, para el reconocimiento de actividades. Por ejemplo, si una persona se aproxima al refrigerador, el resultado del reconocimiento será “abrir refrigerador”, Otro ejemplo, sería que una persona se aproxime a la estufa y la actividad reconocida será “cocinar”.

El modelo logra el reconocimiento de 16 actividades de la vida diaria, listadas a continuación en la Tabla 4.2.

Tabla 4.2. Lista de Actividades Reconocidas

Habitación	Actividad	Nomenclatura
Cocina	Abrir Refrigerador	A1
	Lavar vajilla	A2
	Cocinar	A3
	Salir al jardín	A4
	Preparar alimentos	A5
Comedor	Comer	B1
	Tomar fruta	B2
	Usar microondas	B3
Sala	Hablar por teléfono	C1
	Ver TV	C2
	Leer periodico	C3
	Salir de casa	C4
Dormitorio	Dormir	D1
	Trabajar en PC	D2
	Ejercitarse	D3
	Usar tocador	D4

4.6.1 Medida de proximidad

El modelo propuesto calcula una medida de proximidad, que indica una medida de distancia que existe con respecto a los objetos previamente detectados en cada una de las habitaciones y la posición de las articulaciones, resultado de la estimación de la pose humana para efectuar el reconocimiento de las actividades que realiza una persona.

La métrica de proximidad calculada es la distancia *Manhattan* o longitud *Manhattan* (Krause E. F, 1975) es una métrica en la que la distancia entre dos puntos es la suma de las diferencias absolutas de sus coordenadas. La distancia del *Manhattan*, d_1 , entre dos vectores (p, q) en un espacio vectorial real n-dimensional con un sistema de coordenadas cartesianas fijo es la suma de las longitudes de las proyecciones del segmento de línea entre los puntos sobre el sistema de ejes coordenados. Más formalmente la Ecuación 4.1.

$$d_1(p, q) = \|p - q\|_1 = \sum_{i=1}^n \|p_i - q_i\| \quad (4.1)$$

Donde $p = (p_1, p_2 \dots p_n)$ y $q = (q_1, q_2 \dots q_n)$ son vectores.

En la Figura 4.20, se representan los centros de los objetos previamente reconocidos por el modelo, y se ilustran las articulaciones seleccionadas para realizar el cálculo de la medida de proximidad.

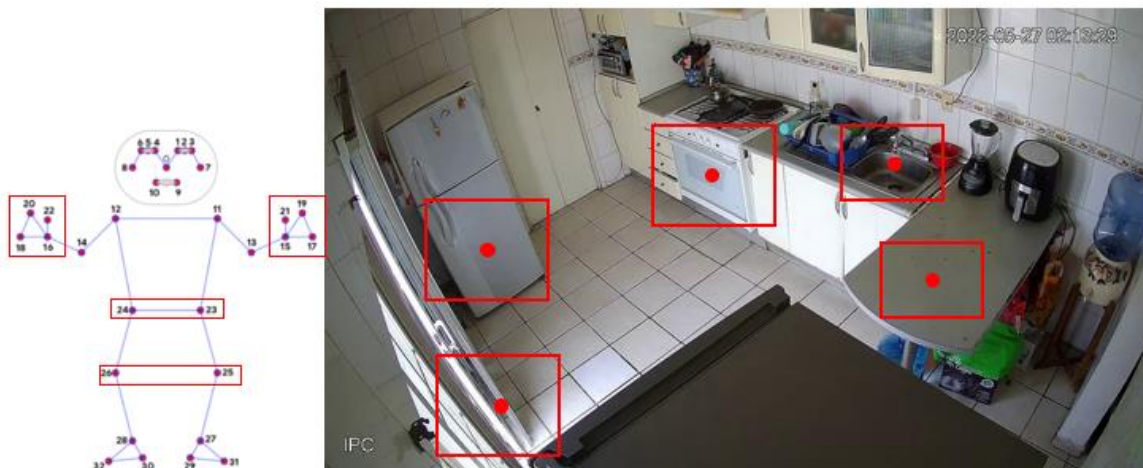


Figura 4.20. Representación de cálculo de medida de proximidad

4.6.2 Secuencia de actividades

El modelo procesa un video y es capaz reconocer una actividad o más de una actividad, que fueron hechas por la persona monitoreada. Es decir, que en ocasiones una persona puede realizar más de una actividad en una misma habitación, lo que da como resultado una secuencia de actividades con el mismo orden en que la persona las realiza.

Por ejemplo, cuando se procesa un video y se reconoce más de una actividad realizada en una misma habitación, el modelo presenta la secuencia de actividades final reconocida en este video y permite almacenar la información de esta secuencia de actividades.

Esto es posible para cualquiera de las habitaciones monitoreadas, lo cual indica que, al terminar de procesar los videos de un día de monitoreo, se tendrá como resultado una secuencia de actividades final, resultado de monitorear las actividades diarias en las cuatro habitaciones de una persona monitoreada.

4.7 Etapa de modelado de rutina

En la etapa de modelado de la rutina se utiliza la información proporcionada por el modelo resultado de las etapas anteriores. Se toma la información de la secuencia de actividades reconocidas en los videos procesados.

Para cada una de las habitaciones, el modelo genera grafos dirigidos, con la información de la secuencia de actividades. Con la creación de grafos dirigidos que se conforma de nodos o estados y transiciones de un estado a otro, se va trazando la rutina de actividades realizadas por la persona que está siendo monitoreada.

Al término de procesar toda la rutina de actividades de una persona, con la duración establecida para la experimentación, en este caso de siete días, el modelo generará un grafo dirigido final, en el que se representarán los patrones de comportamiento de la persona estudiada.

4.7.1 Grafo dirigido

El modelo es entrenado con una rutina diaria y a partir del procesamiento de videos y la información de las secuencias de actividades es capaz de generar grafos dirigidos para cada una de las habitaciones. Donde los nodos o estados representan las actividades y las aristas trazan las transiciones de una actividad a otra (consultar Tabla 4.3). El recorrido del grafo estaría representando información de la secuencia de actividades que se realizan en la rutina de la persona estudiada.

La definición formal de un grafo dirigido (Wasserman, et al., 2013), está definida por un par de conjuntos $G = (V, E)$, donde:

- $V \neq \emptyset$, un conjunto no vacío de objetos simples llamados vértices o nodos.
- $E \subseteq \{(a, b) \in V \times V : a \neq b\}$ es un conjunto de pares ordenados de elementos de V denominados aristas o arcos, donde por definición un arco va del primer nodo (a) al segundo nodo (b) dentro del par.

Un arco $e = (x, y)$ se considera dirigido desde x hacia y . Otra notación válida es $e = \langle x, y \rangle$. En ambos casos, el vértice x cumple un rol de «emisor» y el vértice y uno de «receptor».

Tabla 4.3. Descripción de Grafo dirigido

Origen	Transición	Valor	Destino
Indica estado de origen	Indica el carácter asociado a la transición de un estado actual a un estado siguiente	Representa la frecuencia de ocurrencia del nodo destino	Indica estado de destino

4.8 Etapa de reconocimiento de anomalía

En esta sección del documento se describe el proceso que se lleva a cabo para realizar el reconocimiento de anomalías en el comportamiento de una persona.

En la etapa anterior se ha descrito la fase en la que el modelo genera grafos dirigidos para cada una de las habitaciones, además de un grafo general, resultado de procesar todos los videos de la rutina establecida para la experimentación.

A través de los grafos y el recorrido de los grafos que representan estados de inicio, transiciones y estados finales, el modelo ha creado patrones de comportamiento de la persona monitoreada.

Un patrón de comportamiento en el enfoque de esta investigación está conformado por los recorridos de los grafos creados por el modelo, que contienen información de la secuencia de actividades por habitación y por cada día de monitoreo de la semana.

El patrón de comportamiento, contiene estado de inicio, transiciones y estado final. Además, cuando el modelo va generando estos patrones, asocia un carácter a las transiciones para recorrer el grafo de un estado inicial a un estado final. Como resultado, entonces el modelo genera un patrón de caracteres que representan una secuencia de actividades reconocidas en los videos de una rutina.

El patrón de caracteres será útil para la fase de reconocimiento de anomalías en el comportamiento de una persona, en esta investigación se utilizan máquinas de estado finito como reconocedoras de caracteres.

4.8.1 Máquina de estado finito

El modelo propuesto utiliza máquinas de estado finito (*Finite State Machine*) también llamado autómata finito como reconocedoras, a través de un alfabeto, transiciones, un estado inicial y conjunto de estados finales, se puede obtener un patrón resultado de comenzar en un estado inicial y finalizar con un estado de aceptación o no aceptación.

Se construye una máquina de estados finitos para cada una de las habitaciones y una máquina de estado finito general. En total el modelo implementa cinco máquinas como reconocedoras de patrones. Para su definición se dispone de la información de estados de inicio, transiciones y estado destino de los grafos creados en la etapa anterior.

El patrón de comportamiento contiene información de la secuencia de actividades realizadas por una persona y el modelo se encarga de asociar un carácter al crear los patrones de comportamiento. Estos caracteres forman parte del alfabeto de las máquinas de estado finito para poder ejecutar su función como reconocedoras.

La máquina de estado finito (ver Tabla 4.4), define los estados de aceptación a través del procesamiento de todos los videos que conforman la rutina de la persona monitoreada y permite reconocer si el patrón introducido es aceptado “normal” o rechazado “anormal”.

Tabla 4.4. Definición formal de máquina de estado finito

Notación	Definición
Q	Conjunto finito de estados
Σ	Alfabeto finito
$q_0 \in Q$	Estado inicial
$\delta: Q \times \Sigma \rightarrow Q$	Función de transición
$F \subseteq Q$	Conjunto de estados finales o de aceptación.

Capítulo 5

Validación y experimentación

En este capítulo se presenta la experimentación de cada fase del modelo propuesto, así como las pruebas realizadas para la detección de anomalías. Se inicia con una descripción del ambiente de pruebas que describe el equipo de hardware utilizado para la experimentación, se describe el banco de datos utilizado y sus características. Se detalla el plan de pruebas de la investigación y los resultados obtenidos para evaluar el modelo propuesto.

5.1 Ambiente de pruebas

Las pruebas fueron desarrolladas en un equipo laptop con las siguientes características, fue necesario utilizar el mismo equipo para emular un segundo sistema operativo para poder implementar algunas de las librerías implementadas.

Equipo Laptop

Sistema Operativo: Microsoft Windows 10

Procesador: Intel Core i7

RAM: 12GB

IDE: Spyder 5.3.1

Lenguaje: Python 3.8.8

Equipo Laptop

Sistema Operativo: UBUNTU 16.04.7 LTS

Procesador: Intel Core i7

RAM: 12GB

Librería: Open pose

Librería: YOLO V3

5.2 Banco de datos

El modelo propuesto estudia información del entorno conformado por cuatro habitaciones de un hogar, es por esta razón que no se hace uso de banco de datos de distintas fuentes. Se creó un banco de videos propio con las características en la Tabla 5.1.

Tabla 5.1. Características de banco de datos

Características	Dataset Propio
Formato de video	.mp4
Resolución del video	HD 720 (1280×720), 30 FPS
Duración del video	Aprox. Tres minutos
Actividades	Actividades de la vida diaria
Total de videos	150

5.3 Experimentación

En esta sección del documento se detallará la experimentación realizada para cada una de las fases que conforman el procesamiento de la información del modelo, desde la localización de una persona en el hogar, hasta la construcción de las máquinas de estado finito como reconocedoras de anomalía en el comportamiento de una persona.

5.4 Etapa de localización de una persona

Se utilizaron 3 métodos para la localización de una persona en el hogar, aunque los tres métodos son exitosos y logran su objetivo, se destacan ventajas y desventajas de su implementación, descritas a continuación.

Sensor RFID: Es una tecnología avanzada para la localización de una persona a través de su sensor lector y las tarjetas de radiofrecuencia, sin embargo, mostraron un corto alcance, es decir, el lector tendría que estar posicionado a una distancia aproximada a los 6 cm, lo que a largo plazo se volvería incómodo para la persona monitoreada. Se desea evitar la incomodidad del usuario.

Pulsadores eléctricos: Resultó un método más robusto para el usuario, debido a su bajo costo de producción, el uso de pulsadores fue incómodo de manipular para el usuario, porque es un artefacto que debe aprender a usar. El portar un aparato adicional, suele ser una desventaja para una persona.

Aplicación móvil: Este método obtuvo los mejores resultados, debido a que se usa en un teléfono inteligente, una ventaja es que la persona monitoreada suele estar más familiarizada con su uso, además no tendría que estar portando sensores adicionales, ya que con el mismo teléfono inteligente se puede localizar a la persona y monitorear los sensores acelerómetros y giroscopio utilizados en una etapa posterior.

5.5 Etapa de estimación de la pose humana

Open Pose

Para la estimación de la postura humana, según lo presentado en el estado del arte, una de las librerías con mayor implementación era Open Pose, por esta razón se decidió implementarse para esta investigación.

Se procesaron videos con secuencias de actividades, para obtener la estimación de la postura humana. Sin embargo, los resultados no fueron favorables, los requisitos de hardware GPU que requiere 12 GB de memoria que representan un impacto significativo en el correcto funcionamiento de la librería, especialmente en el coste temporal, o tiempo en el que procesa un video para obtener la estimación de la pose.

Como resultado de la implementación se alcanzó la estimación de la postura humana en el procesamiento de video, pero el tiempo final del procesamiento de video fue extenso para la investigación. Se estima que el tiempo para procesar un solo video de la rutina diaria con duración establecida representaba una duración total de 40 a 50 minutos.

Para no comprometer los objetivos de esta investigación, se descartó la implementación de esta librería, se buscó una alternativa con mejor compromiso en tiempo para resolver la estimación de la pose humana. Los resultados obtenidos de la experimentación se pueden consultar en el **Anexo A**.

MediaPipe Pose

Al implementar la estimación de la pose humana utilizando la librería *MediaPipe Pose*, se obtuvieron mejores resultados en tiempo de procesamiento del video, sin embargo, se tiene que mencionar puntos importantes de la experimentación a través de esta librería.

Comenzando con que el tiempo de procesamiento mejoró considerablemente y permite realizar la estimación con un costo menor en tiempo, el procesamiento del video se lleva a cabo con un tiempo estimado de tres minutos y 35 segundos. Además de procesar videos, la librería permite implementar la estimación de la pose humana en tiempo real, siendo una ventaja importante en esta investigación.

Esta librería tiene como ventaja que proporciona un parámetro nombrado como *visibility*, para cada una de las poses, este indica el nivel de confianza estimado para cada pose calculada, el parámetro es utilizado por el modelo para evitar una falsa estimación de la pose humana. Por ejemplo, si una pose tiene como valor de *visibility* menor a 80% el modelo descarta la estimación de la pose para realizar el reconocimiento de actividad.

A continuación, se presentan casos identificados a través de la experimentación con esta librería, que debe ser considerado para futuras investigaciones en el área de la estimación de la pose humana.

Un primer caso identificado en la estimación de la pose, corresponde a la dificultad de la librería para realizar la estimación de la pose, cuando la cara de la persona no está frente al ángulo de visión de la cámara, esto se debe a que la librería utiliza como método el modelo *BlazeFace* para detectar a una persona y posteriormente realizar el cálculo de la pose humana.

Para contrarrestar esta desventaja, se tenía que voltear a ver previamente a la cámara, o buscar siempre que el rostro de la persona fuera lo más visible posible por el ángulo de la cámara.

Los resultados en Figura 5.1 – Figura 5.3, se muestran ejemplos del caso cuando el rostro de la persona no se encuentra visible frente al ángulo de visión de la cámara.

Caso rostro no visible: usar microondas Figura 5.1 - Figura 5.3



Figura 5.1. Caso rostro no visible: usar microondas



Figura 5.2. Caso rostro no visible: usar microondas



Figura 5.3. Caso rostro no visible: usar microondas

Caso rostro no visible: leer periódico Figura 5.4 - Figura 5.6



Figura 5.4. Caso rostro no visible: leer periódico

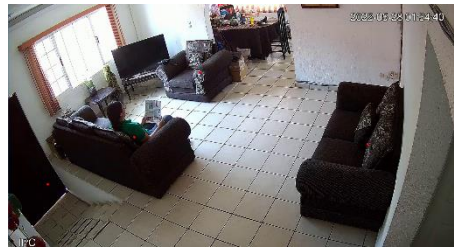


Figura 5.5. Caso rostro no visible: leer periódico



Figura 5.6. Caso rostro no visible: leer periódico

Caso rostro no visible: salir de casa Figura 5.7 - Figura 5.10



Figura 5.7. Caso rostro no visible: salir de casa



Figura 5.8. Caso rostro no visible: salir de casa



Figura 5.9. Caso rostro no visible: salir de casa



Figura 5.10. Caso rostro no visible: salir de casa

Un caso adicional que fue identificado, corresponde a la pérdida de estimación de la pose humana si existe un objeto que ocluya parcialmente la postura de una persona. Este caso se presentó en mayor ocasión en tomar fruta, preparar alimento y comer.

Caso obstrucción de la postura: tomar fruta Figura 5.11 - Figura 5.13

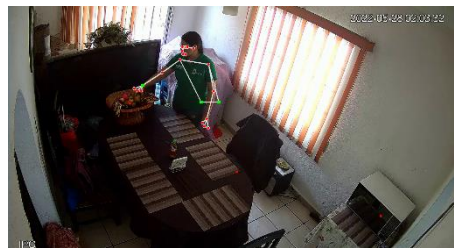


Figura 5.11. Caso rostro no visible: tomar fruta

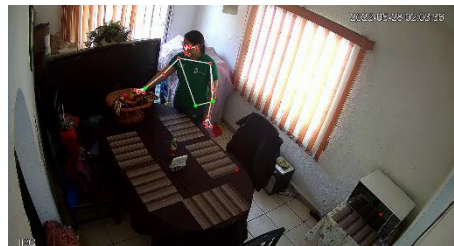


Figura 5.12. Caso rostro no visible: tomar fruta



Figura 5.13. Caso rostro no visible: tomar fruta

Caso obstrucción de la postura: preparar alimento Figura 5.14 - Figura 5.16



Figura 5.14. Caso obstrucción de la postura: preparar alimento

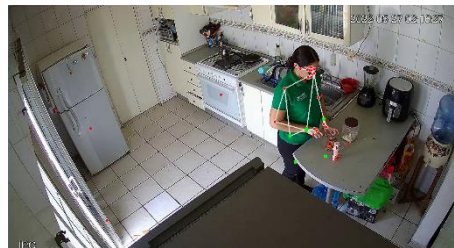


Figura 5.15. Caso obstrucción de la postura: preparar alimento

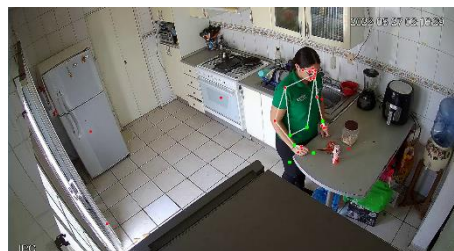


Figura 5.16. Caso obstrucción de la postura: preparar alimento

Caso obstrucción de la postura: comer Figura 5.17 - Figura 5.19

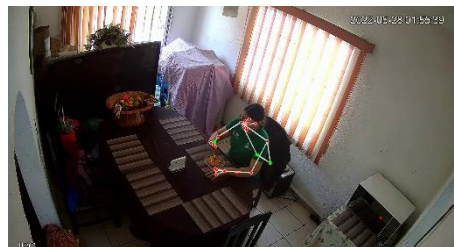


Figura 5.17. Caso obstrucción de la postura: comer

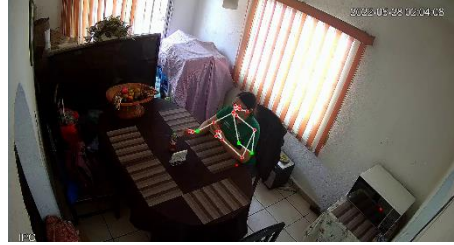


Figura 5.18. Caso obstrucción de la postura: comer

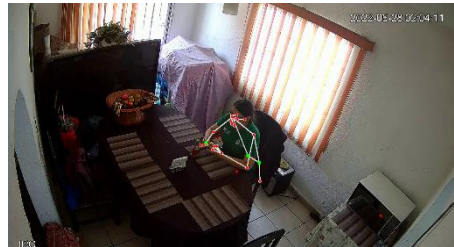


Figura 5.19. Caso obstrucción de la postura: comer

Un caso más corresponde a la falsa estimación de la pose humana, es decir, que la librería estima la postura humana de forma errónea sin que se trate de la postura real de una persona. Se detectó estos casos en ciertas texturas presentes en el video, en este tipo de casos se descarta la estimación de la pose utilizando el parámetro *visibility*. La identificación de estos casos permitió evitar el reconocimiento de actividades de manera equivocada.

Caso falsa estimación de la postura: Figura 5.20 - Figura 5.23

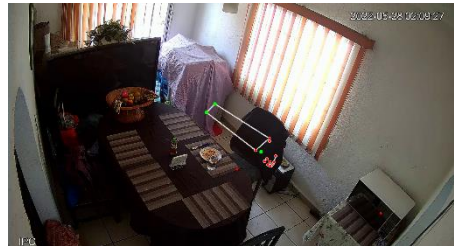


Figura 5.20. Caso falsa estimación de la postura



Figura 5.21. Caso falsa estimación de la postura



Figura 5.22. Caso falsa estimación de la postura



Figura 5.23. Caso falsa estimación de la postura

Los resultados obtenidos se muestran en las siguientes tablas 5.2 – 5.5 se detalla el porcentaje de la métrica *accuracy* (exactitud) del reconocimiento de las 16 actividades. El cálculo se realizó en relación a una muestra de 100 *frames* obtenidos de los videos de grabación de las actividades.

Tabla 5.2 Habitación: Cocina

Tabla 5.2. Resultados de reconocimiento de actividad: Cocina

Actividad	<i>Accuracy</i> (Exactitud)
Abrir refrigerador	91%
Lavar vajilla	92%
Cocinar	95%
Salir al jardín	92%
Preparar alimentos	95%

Tabla 5.3 Habitación: Comedor

Tabla 5.3. Resultados de reconocimiento de actividad: Comedor

Actividad	<i>Accuracy</i> (Exactitud)
Comer	91%
Tomar fruta	92%
Usar microondas	89%

Tabla 5.4 Habitación: Sala

Tabla 5.4. Resultados de reconocimiento de actividad: Sala

Actividad	Accuracy (Exactitud)
Hablar por teléfono	90%
Ver TV	92%
Leer periódico	89%
Salir de casa	85%

Tabla 5.5 Habitación: Dormitorio

Tabla 5.5. Resultados de reconocimiento de actividad: Dormitorio

Actividad	Accuracy (Exactitud)
Dormir	92%
Trabajar en PC	91%
Ejercitarse	92%
Usar tocador	95%

Los resultados de la experimentación muestran que las actividades reconocidas correctamente fueron: cocinar con 95% de exactitud, preparar alimentos 95%, usar tocador 95%. Las actividades que presentan menor porcentaje de exactitud de reconocimiento resultaron: salir de casa con 85%, usar microondas 89% y leer periódico 89%. Aunque se obtuvo resultados favorables en el reconocimiento de actividades, las actividades salir de casa, usar microondas y leer periódico presentan un porcentaje de exactitud menor, esto a causa del método utilizado para la estimación de la pose humana. Las pruebas muestran que debido a que el rostro de la persona no es visible dentro del rango de visión de la cámara y la metodología de *Mediapipe pose* realiza la detección de rostro para la estimación de las partes del cuerpo, esta circunstancia fue uno de los factores que influye en los resultados de las pruebas realizadas, así como las pérdidas de la pose humana cuando se ocluye parcialmente el cuerpo de una persona y los casos donde se presentó la falsa estimación de la pose.

5.6 Etapa de reconocimiento de objetos

En esta etapa se experimenta con la implementación de la Liberia YOLO V3, los resultados son favorables, se describe el tiempo de predicción de los objetos de cada habitación y además la confianza de predicción de cada objeto. En seguida se presentan los resultados con más detalle en Figura 5.24 - Figura 5.27.

Tabla 5.6 Habitación: Cocina**Tiempo de predicción:** 25.9 segundos.

Tabla 5.6. Resultados de detección de objeto: Cocina

Objeto	Confianza de predicción
Refrigerador	99%
Horno estufa	94%
Fregadero	60%

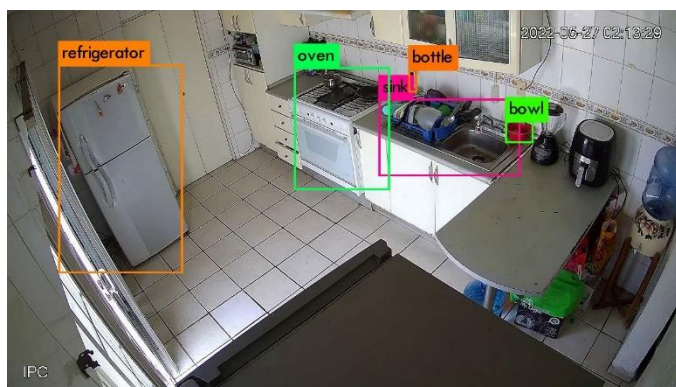


Figura 5.24. Detección de objetos: Cocina

Tabla 5.7 Habitación: Comedor**Tiempo de predicción:** 26.8 segundos.

Tabla 5.7. Resultados de detección de objeto: Comedor

Objeto	Confianza de predicción
Mesa de comedor	84%
Silla	87%

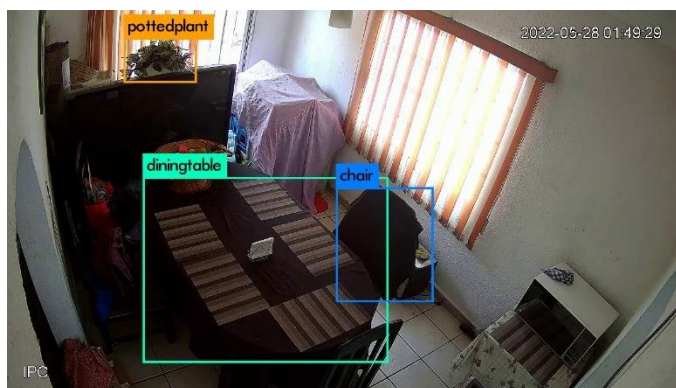


Figura 5.25. Detección de objetos: Comedor

Tabla 5.8 Habitación: Sala**Tiempo de predicción:** 25.9 segundos.

Tabla 5.8. Resultados de detección de objeto: Sala

Objeto	Confianza de predicción
Televisión	77%
Sofá	95%
Sofá	90%

En la Figura 5.26, se obtiene como resultado la clasificación errónea de un sofá como *chair* que se traduce a silla, en este tipo de casos específicos donde la clasificación no fue correcta y el porcentaje de confianza de predicción es menor, se procede a ignorar la detección de objetos clasificados erróneamente y se realiza un etiquetado manual de la zona.

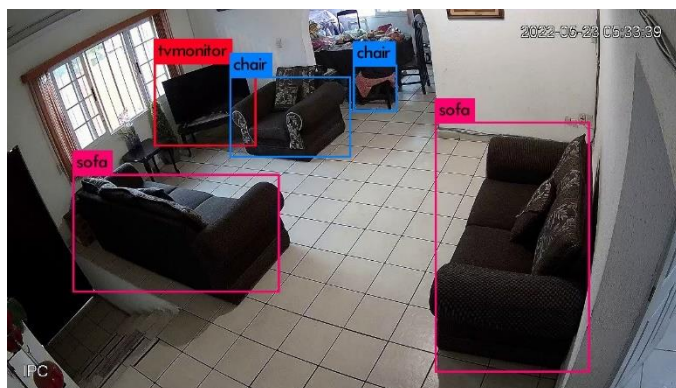


Figura 5.26. Detección de objetos: Sala

Tabla 5.9 Habitación: Dormitorio**Tiempo de predicción:** 25.2 segundos.

Tabla 5.9. Resultados de detección de objeto: Dormitorio

Objeto	Confianza de predicción
Cama	97%
Laptop	97%
Silla	100%

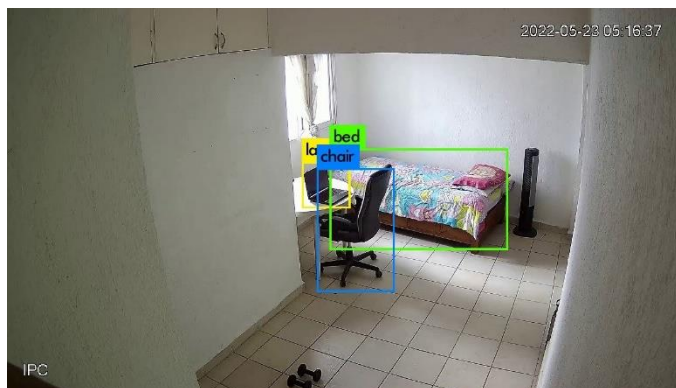


Figura 5.27. Detección de objetos: Dormitorio

En los casos específicos donde el objeto no era visible por completo en el área de visión de las cámaras y no fue detectado por el modelo YOLOv3, se procedió a realizar un etiquetado manual de la zona de interés Figura 5.28 - Figura 5.31, esos casos son enlistados en Tabla 5.10.

Tabla 5.10 Etiquetado manual de objetos

Tabla 5.10. Etiquetado manual de objetos

Habitación	Etiquetado manual del objeto
Cocina	Acceso al jardín, mesa
Comedor	Frutero, horno de microondas
Sala	Puerta
Dormitorio	Tocador, área de ejercicio



Figura 5.28. Etiquetado manual de objetos: Cocina



Figura 5.29. Etiquetado manual de objetos: Comedor



Figura 5.30. Etiquetado manual de objetos: Sala



Figura 5.31. Etiquetado manual de objetos: Dormitorio

5.7 Etapa de recolección de datos de sensores

Como parte de la experimentación realizada, se presenta un experimento adicional al monitoreo de sensores. El experimento se realizó con 10 personas para validar que la fase de monitoreo de sensor se realizara correctamente.

Se indicó a cada uno de los participantes cómo utilizar el teléfono inteligente y dónde portarlo, también se les indicó qué actividades deberían realizar. Se monitoreó 6 actividades, se les pidió realizaran las actividades como normalmente solían realizarlas en su vida cotidiana. Las actividades seleccionadas para el experimento fueron ambulatorias (caminar, trotar, ponerse de pie, subir, bajar y sentarse).

Se entrenó una red neuronal profunda (DNN), para reconocer el tipo de movimiento, en función de un conjunto de datos del sensor acelerómetro de un dispositivo móvil portado en la cintura de una persona. El conjunto de datos utilizado fue WISDM (Jennifer R, et al., 2010) consultar Tabla 5.11, para el entrenamiento de la red con datos de sensores multidimensionales divididos en el tiempo.

Posteriormente, se realizó un conjunto de datos propio detallado en la Tabla 5.14 para entrenar la misma red neuronal profunda (DNN) y comparar los resultados obtenidos. Esto permite validar la recopilación de los datos obtenidos del sensor, además de valorar los resultados de clasificación de la red.

Tabla 5.11 Banco de datos WISDM

Tabla 5.11. Descripción banco de datos WISDM

Participantes	Actividades	Sensor
33	Caminar, trotar, subir, bajar, sentarse, ponerse de pie	Acelerómetro ejes (x,y,z)

Normalización y preprocesamiento de los datos

Se requiere normalizar las características de los datos de entrenamiento. Existen varias formas de normalizar las características, para este experimento se normaliza los 3 ejes del sensor (valores entre 0 y 1) y se aplica un redondeo a las características.

- $x\text{-axis } i = x\text{-axis } i / \text{máximo } (x\text{-axis})$
- $y\text{-axis } i = y\text{-axis } i / \text{máximo } (y\text{-axis})$
- $z\text{-axis } i = z\text{-axis } i / \text{máximo } (z\text{-axis})$

Los datos contenidos en el marco de datos requieren una estructura para ingresar a la red neuronal. Por lo tanto, se debe de adaptar las características al modelo. Teniendo en cuenta la frecuencia de muestreo de 20 Hz, esto equivale a intervalos de tiempo de 4 segundos (cálculo: $0.05 * 80 = 4$). Además de remodelar los datos, la función también separará las

características (aceleración x , aceleración y , aceleración z) y las etiquetas de actividad asociada.

Modelo de red neuronal profunda

Se crea una red neuronal con 3 capas ocultas de 100 nodos, consultar Tabla 5.12, la primera capa en la red neuronal remodelará los datos. Las dos últimas capas volverán a aplanar los datos y luego ejecutarán una función de activación *softmax* para calcular la probabilidad de cada clase. A continuación, se describe la red neuronal.

Tabla 5.12. Modelo de red neuronal profunda

Layer (type)	Output Shape	Param #
reshape (Reshape)	(None, 80, 3)	0
dense (Dense)	(None, 80, 100)	400
dense_1 (Dense)	(None, 80, 100)	10100
dense_2 (Dense)	(None, 80, 100)	10100
flatten (Flatten)	(None, 8000)	0
dense_3 (Dense)	(None, 6)	48006
Total params:	68,606	
Trainable params:	68,606	
Non-trainable params:	0	

Los hiperparámetros utilizados para el entrenamiento son: un lote de 400 registros y se entrenará el modelo durante 50 épocas. Para el entrenamiento de modelos, se usará una división 80:20 para separar los datos de entrenamiento y los datos de validación del conjunto de datos WISDM (Jennifer R, et al., 2010).

Resultado del conjunto de datos WISDM

El rendimiento de la DNN con el conjunto de datos WISDM (Jennifer R, et al., 2010) logra la mejor predicción. Se obtiene una precisión de validación de aproximadamente el 74% en la Figura 5.32 se detalla la los resultados de la matriz de confusión. Esto podría mejorar, con más ajustes de hiperparámetros y especialmente con un diseño de red neuronal modificado, consultar resultados Tabla 5.13.

Accuracy on test data: 0.74

Loss on test data: 1.99

Tabla 5.13. Resultados de WISDM

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
bajar	0.62	0.79	0.69	1866
trotar	0.97	0.98	0.97	6426
sentarse	1.00	1.00	1.00	1048
ponerse de pie	0.99	1.00	1.00	834
subir	0.76	0.68	0.72	2339
caminar	0.93	0.89	0.91	8062
<i>accuracy</i>			0.90	20575
<i>macro avg</i>	0.88	0.89	0.88	20575
<i>weighted avg</i>	0.90	0.90	0.90	20575

Matriz de confusión WISDM

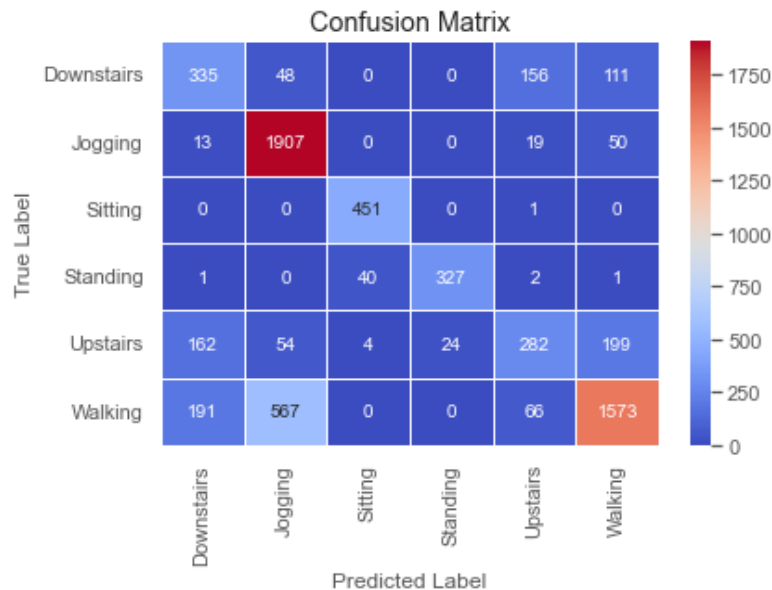


Figura 5.32. Matriz de confusión: WISDM

Como se puede ver en la Figura 5.32, la precisión del modelo logra una mejor puntuación para predecir, trotar, sentarse, pararse y caminar. Sin embargo, el rendimiento para distinguir las actividades subir y bajar, es bastante insatisfactorio. El modelo tiene problemas para identificar claramente las actividades.

Tabla 5.14 Banco de datos propio

Tabla 5.14. Descripción de banco de datos propio

Participantes	Actividades	Sensor
10	Caminar, trotar, subir, bajar, sentarse, ponerse de pie	Acelerómetro ejes (x,y,z)

Resultado del conjunto de datos propio

El experimento consiste en entrenar la misma red neuronal profunda utilizando el conjunto de datos propio de esta investigación y analizar los resultados de la matriz de confusión Figura 5.33. Se realizó la recopilación de datos del sensor acelerómetro con 5 participantes para las seis actividades (caminar, trotar, subir, bajar, sentarse, ponerse de pie) se implementó el preprocesamiento de los datos y los resultados ver Tabla 5.15.

- **Accuracy on test data:** 0.64
- **Loss on test data:** 2.25

Tabla 5.15. Resultados de conjunto de datos propio

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
bajar	0.66	0.87	0.75	285
trotar	0.99	0.98	0.99	1029
sentarse	0.99	1.00	1.00	114
ponerse de pie	1.00	1.00	1.00	108
subir	0.95	0.93	0.94	281
caminar	0.97	0.92	0.95	1395
<i>accuracy</i>			0.94	3212
<i>macro avg</i>	0.93	0.95	0.94	3212
<i>weighted avg</i>	0.95	0.94	0.94	3212

Matriz de confusión conjunto propio

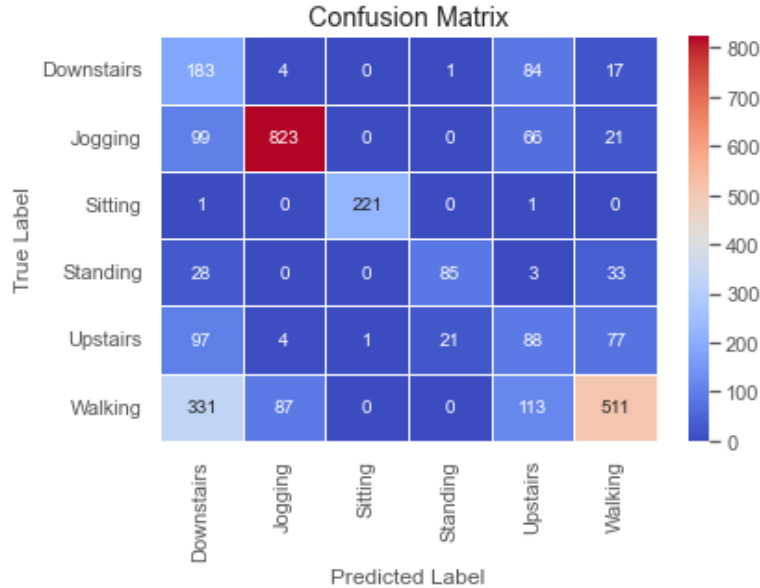


Figura 5.33. Matriz de confusión: conjunto propio

El experimento realizado permite validar la recopilación de los datos del sensor para esta investigación, en los resultados se puede observar que el conjunto de datos propio obtuvo una precisión menor 64% en comparación con el conjunto de datos WISDM (Jennifer R, et al., 2010) 74%, sin embargo, para la clasificación de las actividades subir 95% y bajar 66% se visualizan en la Figura 5.33 mejores resultados que en el conjunto WISDM (Jennifer R, et al., 2010).

Todavía existe un gran potencial para mejorar el modelo, por ejemplo, mediante el uso de diseños de redes neuronales más avanzados como las redes neuronales convolucionales (CNN) o la memoria a largo plazo (LSTM).

Los resultados indican que la manera en la que se está monitoreando a una persona a través de los sensores es a través de la frecuencia de muestreo y posición adecuada.

La validación descrita en los párrafos anteriores fue realizada con el propósito de comprobar que el método es correcto, sin embargo, como parte del modelo, el monitoreo de sensores se utiliza únicamente para monitorear a la persona dentro de los puntos ciegos de las cámaras de monitoreo del entorno donde se lleva a cabo la experimentación.

Los resultados de implementar el monitoreo de los sensores acelerómetro y giroscopio a una frecuencia de 60 Hz, en el bolsillo del pantalón de la persona, se describen a continuación. Se muestra las señales de los sensores con el fenómeno de espectro de ruido y después los resultados de aplicar un filtro de ruido de la señal para cada uno de los sensores.

En la Figura 5.34, se puede observar la lectura de los ejes del sensor acelerómetro con el fenómeno de espectro de ruido en las señales.

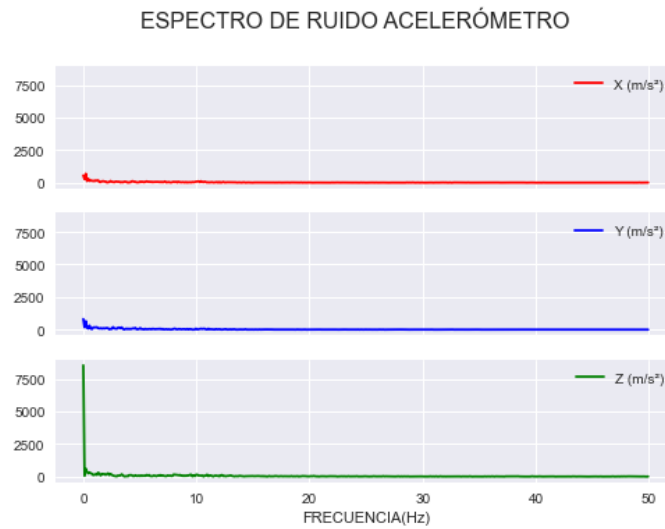


Figura 5.34. Representación de espectro de ruido en acelerómetro

En la Figura 5.35, se puede observar la lectura de los ejes del acelerómetro después de atenuar las señales del sensor.

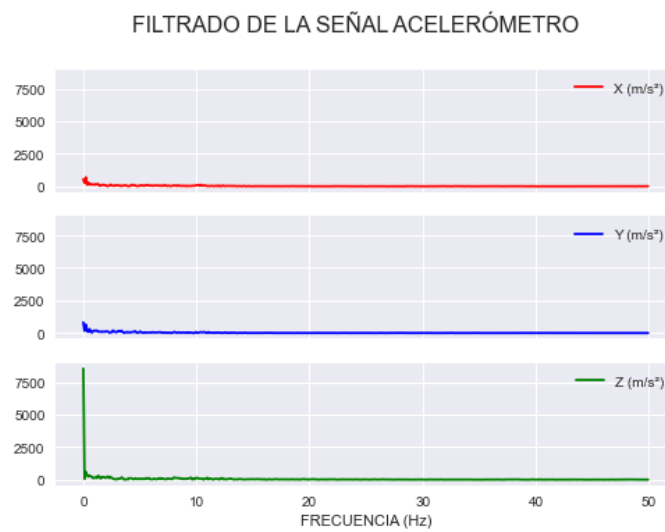


Figura 5.35. Representación de filtrado de la señal del acelerómetro

En la Figura 5.36, se puede observar la lectura de los ejes del sensor giroscopio con el fenómeno de espectro de ruido en las señales.

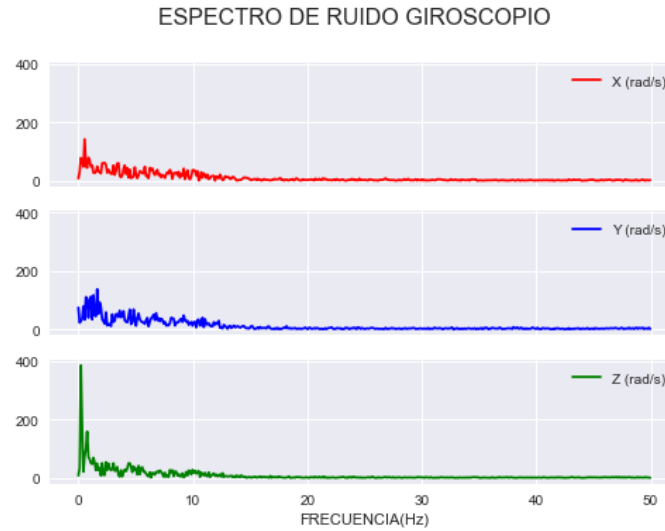


Figura 5.36. Representación de espectro de ruido en giroscopio

En la Figura 5.37, se puede observar la lectura de los ejes del giroscopio después de atenuar las señales del sensor.

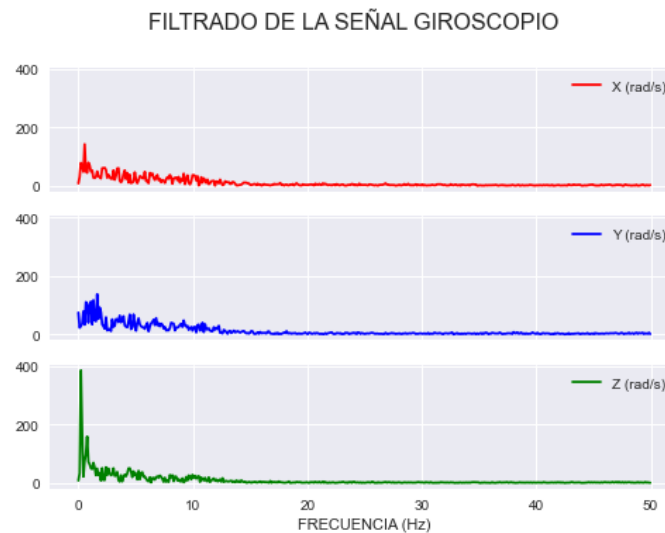


Figura 5.37. Representación de filtrado de la señal del giroscopio

5.8 Etapa de reconocimiento de actividad

En esta etapa se realiza el proceso de la estimación de la pose humana y el cálculo de proximidad para el reconocimiento de actividades. En la experimentación previamente se ha estimado la postura humana y el cálculo de una medida de proximidad.

Se evalúa la distancia Manhattan aplicada a la estimación de la postura humana que corresponde a las articulaciones con número (15, 16, 23, 24, 25, 26), calculadas por *MediaPipe Pose*, (para identificar las articulaciones consultar Figura 4.13).

Las articulaciones seleccionadas en la estimación de la pose de la persona son resaltadas en color verde en la Figura 5.38, así como las coordenadas (x,y) de la posición del centro del objeto previamente detectado, que se supone la persona se encontrará interactuando de cerca.



Figura 5.38. Cálculo de medida de proximidad

Los resultados del reconocimiento de 16 actividades de la vida diaria se presentan en la Tabla 5.16, donde se muestra la actividad reconocida por el modelo y se detallan observaciones sobresalientes.

Tabla 5.16. Reconocimiento de actividades



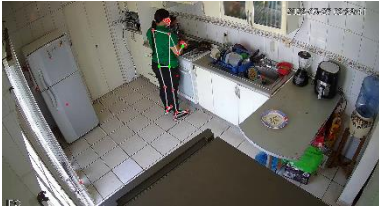


Actividad	Imagen	Observaciones
Abrir Refrigerador		Se estima la pose desde que la persona entra en la escena. Al verse completamente la postura, el modelo no presenta dificultad para realizar el reconocimiento de la actividad asociada.
Lavar vajilla		Se estima la pose sin dificultad, sin embargo, se puede observar que las piernas de la persona se encuentran ligeramente obstruidas, por lo tanto, el gráfico de la pose suele presentar una inclinación en el área de las piernas
Cocinar		Se estima la pose desde que la persona entra en la escena. Al verse completamente la postura, el modelo no presenta dificultad para realizar el reconocimiento de la actividad asociada.
Salir al jardín		Se estima la pose, sin embargo, se puede observar que las piernas de la persona no se encuentran visibles, como resultado, el gráfico de la pose no se presenta completo.
Preparar alimentos		Se estima la pose, sin embargo, se puede observar que las piernas de la persona no se encuentran visibles, como resultado, el gráfico de la pose no se presenta completo.

Tabla 5.16. Reconocimiento de actividades (Continuación)

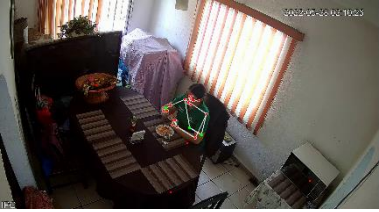
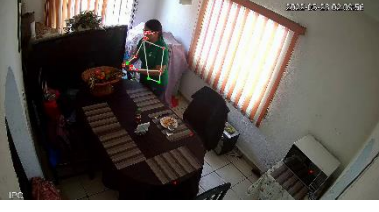
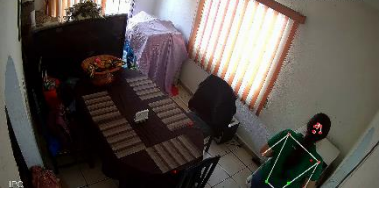
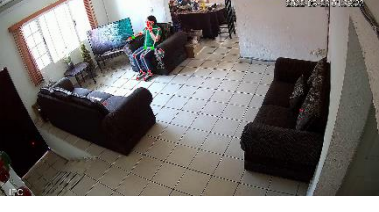
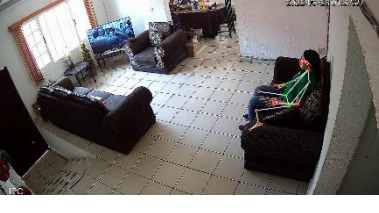
Actividad	Imagen	Observaciones
Comer		Se estima la pose, sin embargo, se puede observar que las piernas de la persona no se encuentran visibles, como resultado, el gráfico de la pose no se presenta completo.
Tomar fruta		Se estima la pose, sin embargo, se puede observar que las piernas de la persona no se encuentran visibles, como resultado, el gráfico de la pose no se presenta completo.
Usar microondas		Se estima la pose, sin embargo, se puede observar que las piernas de la persona no se encuentran visibles, por lo tanto, el gráfico de la pose no se presenta completo. El reconocimiento de la actividad presenta una mayor demora para el sistema.
Hablar por teléfono		La librería estima la pose desde que la persona entra en la escena. Al verse completamente la postura, el modelo no presenta dificultad para realizar el reconocimiento de la actividad asociada.
Ver TV		Se estima la pose sin dificultad, sin embargo, se puede observar que las piernas de la persona se encuentran ligeramente obstruidas, como resultado, el gráfico de la pose suele presentar una inclinación en el área de las piernas

Tabla 5.16. Reconocimiento de actividades (Continuación)

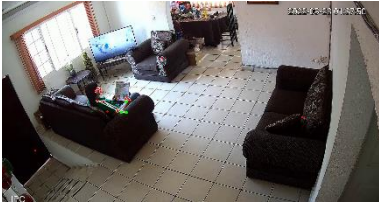



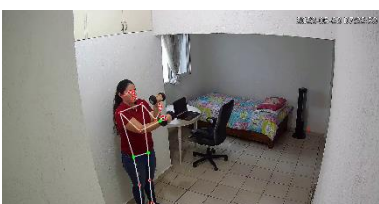

Actividad	Imagen	Observaciones
Leer periódico		<p>La estimación de la pose se realiza con dificultad porque el rostro y las piernas de la persona no es visible en su totalidad. El reconocimiento de la actividad presenta una mayor demora para el sistema.</p>
Salir de casa		<p>La estimación de la pose se realiza con dificultad porque el rostro y las piernas de la persona se contrastan en la textura de la puerta del hogar. El reconocimiento de la actividad presenta una mayor demora para el sistema.</p>
Dormir		<p>Se estima la pose desde que la persona entra en la escena. Al verse completamente la postura, el modelo no presenta dificultad para realizar el reconocimiento de la actividad asociada.</p>
Trabajar en PC		<p>Se estima la pose desde que la persona entra en la escena. Al verse completamente la postura, el modelo no presenta dificultad para realizar el reconocimiento de la actividad asociada.</p>
Ejercitarse		<p>Se estima la pose desde que la persona entra en la escena. Al verse completamente la postura, el modelo no presenta dificultad para realizar el reconocimiento de la actividad asociada.</p>

Tabla 5.16. Reconocimiento de actividades (Continuación)

Actividad	Imagen	Observaciones
Usar tocador		Se estima la pose sin dificultad, sin embargo, se puede observar que las piernas de la persona se encuentran ligeramente obstruidas, por lo tanto, el gráfico de la pose suele presentar una pérdida en el área de las piernas

Secuencia de actividades

En los resultados del procesamiento de un video por parte del modelo se obtiene el reconocimiento de una actividad o más de una actividad presente en el video, dando lugar a una secuencia de actividades. Esta secuencia se determina cuando una persona entra a una habitación, por ejemplo Figura 5.39, una persona entra a la cocina y realiza en el siguiente orden “Abrir refrigerador”, “Cocinar” y “Preparar alimentos”, como resultado el modelo reconoce la secuencia de actividades: A1, A3, A5 (consultar **¡Error! No se encuentra el origen de la referencia.** para conocer los identificadores de las actividades).



Figura 5.39. Secuencia de actividades

5.9 Etapa de modelado de la rutina

El enfoque que tiene esta etapa es procesar todos los videos de una rutina, realizar el reconocimiento de actividades en un video, utilizar la información de las secuencias de actividades y convertir la información para generar grafos dirigidos. Como se mencionó anteriormente, el modelo genera grafos dirigidos de cada una de las cuatro habitaciones y un grafo general resultado de procesar todos los videos de la rutina para la experimentación del modelo.

Los grafos generados se presentan en las Figura 5.40 - Figura 5.44, además de definir la información de estado origen, transición, valor y estado destino en las tablas Tabla 5.17 - Tabla 5.21.

Grafo: Cocina

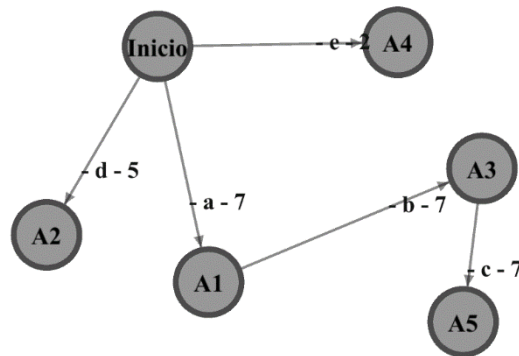


Figura 5.40. Grafo: Cocina

Tabla 5.17. Tabla de transiciones grafo: Cocina

Origen	Transición	Valor	Destino
Inicio	a	7	A1
Inicio	d	5	A2
Inicio	e	2	A4
A1	b	7	A3
A3	c	7	A5

Grafo: Comedor

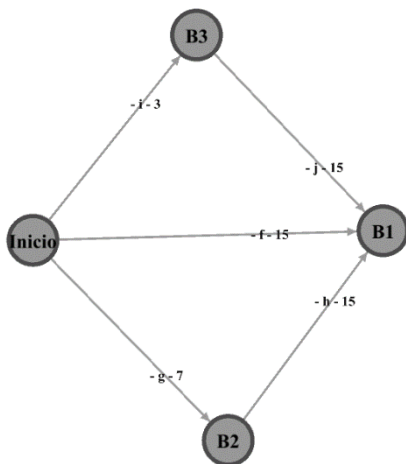


Figura 5.41. Grafo: Comedor

Tabla 5.18. Tabla de transiciones grafo: Comedor

Origen	Transición	Valor	Destino
Inicio	f	15	B1
Inicio	g	7	B2
Inicio	i	3	B3
B2	h	15	B1
B3	j	15	B1

Grafo: Sala

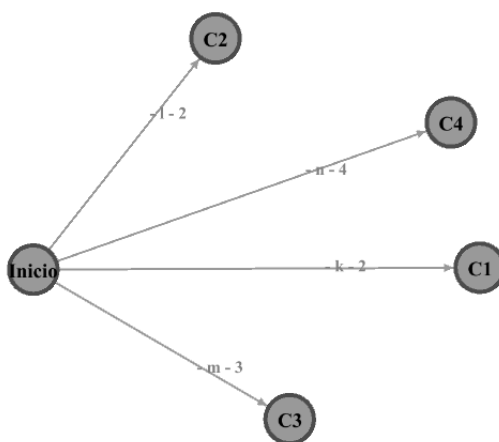


Figura 5.42. Grafo: Sala

Tabla 5.19. Tabla de transiciones grafo: Sala

Origen	Transición	Valor	Destino
Inicio	k	2	C1
Inicio	l	2	C2
Inicio	m	3	C3
Inicio	n	4	C4

Grafo: Dormitorio

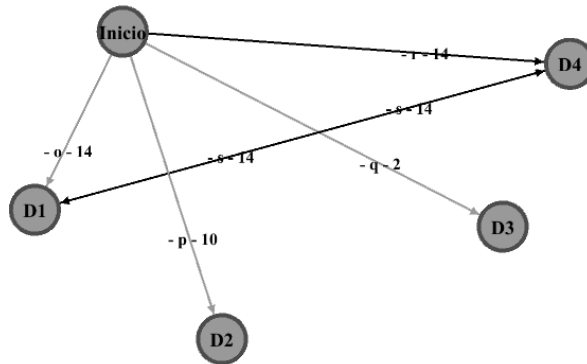


Figura 5.43. Grafo: Dormitorio

Tabla 5.20. Tabla de transiciones grafo: Dormitorio

Origen	Transición	Valor	Destino
Inicio	o	14	D1
Inicio	p	10	D2
Inicio	q	2	D3
Inicio	r	14	D4
D1	s	14	D4
D4	s	14	D1

Grafo general

En este grafo general (ver Figura 5.44 pág. 72), se representan todas las secuencias de actividades que realiza una persona en su rutina diaria, se observan las aristas que contienen información acerca de las transiciones de un estado a otro estado, además del valor numérico de las aristas que representan la frecuencia en la que se realiza la actividad asociada a los estados destino del grafo.

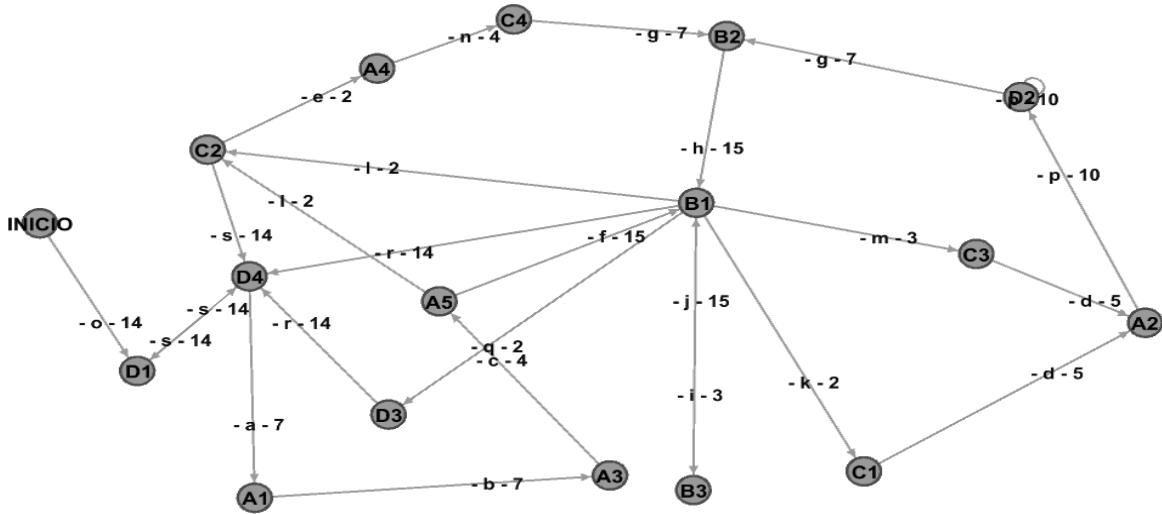


Figura 5.44. Grafo: General

Transiciones grafo general

La información de la tabla de transiciones Tabla 5.21, será utilizada para definir la máquina de estados finito general del modelo, misma que permitirá realizar la clasificación de patrones de comportamiento de la persona monitoreada.

Tabla 5.21. Tabla de transiciones grafo: General

Origen	Transición	Valor	Destino
D4	a	7	A1
C1	d	5	A2
C3	d	5	A3
A1	b	7	A3
C2	e	2	A4
A3	c	2	A5
A5	f	4	B1
B2	h	15	B1
B3	j	15	B1
C4	g	15	B2
D2	g	7	B2
B1	i	7	B3
B1	k	3	C1
B1	l	2	C2
B1	m	2	C3
A4	n	3	C4
Inicio	o	14	D1

Tabla 5.21. Tabla de transiciones grafo: General (Continuación)

Origen	Transición	Valor	Destino
D4	s	14	D1
D2	p	10	D2
A2	p	10	D2
B1	q	2	D3
B1	r	14	D4
D1	s	14	D4
A5	l	2	C2
C2	r	14	D4

5.10 Etapa de reconocimiento de anomalía

En esta etapa se presentan los resultados de la definición de las máquinas de estado finito que implementa el modelo para hacer el reconocimiento de anomalías en el comportamiento de una persona.

Para cada una de las habitaciones se utiliza una máquina de estado finito que define estados finitos, alfabeto, un estado de inicio y un conjunto de estados finales. Además, se presenta una tabla con las transiciones para cada máquina utilizada.

Además, se presenta una máquina de estado finito especificada como general en la Tabla 5.30 que incluye la información del grafo dirigido general (ver Tabla 5.21) que representa que se han procesado todos los videos que conforman la rutina diaria de una persona.

Las máquinas de estado finito se describen en esta sección en las siguientes Tabla 5.22 – Tabla 5.31.

Máquina de estado finito: Cocina

Tabla 5.22. Definición formal de máquina de estado finito: Cocina

Q	Inicio, A1, A2, A3, A4, A5
Σ	a, b, c, d, e
$q_0 \in Q$	Inicio
$\mathcal{F} \subseteq Q$	A2, A4, A5

Tabla 5.23. Tabla de transiciones máquina de estado finito: Cocina

$q \in Q$	$\sigma \in \Sigma$	$\delta(q, \sigma) \in Q$
Inicio	a	A1
Inicio	d	A2
Inicio	e	A4
A1	b	A3
A3	c	A5

Máquina de estado finito: Comedor

Tabla 5.24. Definición formal de máquina de estado finito: Comedor

Q	Inicio, B1, B2, B3
Σ	f, g, h, i, j
$q_0 \in Q$	Inicio
$\mathcal{F} \subseteq Q$	B1

Tabla 5.25. Tabla de transiciones máquina de estado finito: Comedor

$q \in Q$	$\sigma \in \Sigma$	$\delta(q, \sigma) \in Q$
Inicio	f	B1
Inicio	g	B2
Inicio	i	B3
B2	h	B1
B3	j	B1

Máquina de estado finito: Sala

Tabla 5.26. Definición formal de máquina de estado finito: Sala

Q	Inicio, C1, C2, C3, C4
Σ	k, l, m, n
$q_0 \in Q$	Inicio
$\mathcal{F} \subseteq Q$	C1, C2, C3, C4

Tabla 5.27. Tabla de transiciones máquina de estado finito: Sala

$q \in Q$	$\sigma \in \Sigma$	$\delta(q, \sigma) \in Q$
Inicio	k	C1
Inicio	l	C2
Inicio	m	C3
Inicio	n	C4

Máquina de estado finito: Dormitorio

Tabla 5.28. Definición formal de máquina de estado finito: Dormitorio

Q	Inicio, D1, D2, D3, D4
Σ	o, p, q, r, s
$q_0 \in Q$	Inicio
$F \subseteq Q$	D1, D2, D3, D4

Tabla 5.29. Tabla de transiciones máquina de estado finito: Dormitorio

$q \in Q$	$\sigma \in \Sigma$	$\delta(q, \sigma) \in Q$
Inicio	o	D1
Inicio	p	D2
Inicio	q	D3
Inicio	r	D4
D1	s	D4
D4	s	D1

Máquina de estado finito: General

La definición de esta máquina de estado finito incluye todos los posibles estados que en el enfoque de esta investigación resultan ser las actividades diarias de la rutina de una persona monitoreada (ver

Tabla 5.30) que contiene un alfabeto que representa los caracteres de un patrón de comportamiento, un estado de inicio que se traduce a la actividad con la que suele iniciar su rutina una persona y un estado de aceptación que se traduce a la actividad con la que suele terminar la rutina la persona monitoreada. Después se presenta en la Tabla 5.31 de transiciones que se debe cumplir para transitar de un estado actual a un estado siguiente.

Tabla 5.30. Definición formal de máquina de estado finito: General

Q	A1, A2, A3, A4', A5, B1, B2, B3, C1, C2, C3, C4, D1, D2, D3, D4, Inicio
Σ	a, b, c, d, k, l, f, g, h, i, j, m, n, e, o, p, q, r, s
$q_0 \in Q$	Inicio
$\mathcal{F} \subseteq Q$	D1

Tabla 5.31. Tabla de transiciones máquina de estado finito: General

$q \in Q$	$\sigma \in \Sigma$	$\delta(q, \sigma) \in Q$
D4	a	A1
C1	d	A2
C3	d	A3
A1	b	A3
C2	e	A4
A3	c	A5
A5	f	B1
B2	h	B1
B3	j	B1
C4	g	B2
D2	g	B2
B1	i	B3
B1	k	C1
B1	l	C2
B1	m	C3
A4	n	C4
Inicio	o	D1
D4	s	D1
D2	p	D2
A2	p	D2
B1	q	D3
B1	r	D4
D1	s	D4

5.11 Plan de pruebas

El objetivo del plan de pruebas es validar el correcto funcionamiento del modelo y comprobar que se alcanzaron los objetivos de esta investigación.

Para realizar el plan de pruebas primero se planifica y organiza la rutina diaria que será realizada por la persona monitoreada, posteriormente se realizaron pruebas con la rutina diaria actual de la persona, la cual se va a monitorear para realizar el reconocimiento de anomalías en comparación con la primera rutina utilizada para la definición del modelo.

En esta sección se detalla la rutina diaria utilizada para la fase de entrenamiento del modelo y también se proporciona información de las pruebas para validar que el modelo reconoce anomalías en el comportamiento de una persona, en caso de que existieran.

5.11.1 Entrenamiento del modelo

Para seleccionar la duración del periodo de la rutina diaria de una persona, se realizó un análisis de la duración de monitoreo de algunas de las investigaciones del estado del arte. En esta investigación se decide comenzar por una duración de siete días de la semana, tiempo suficiente para representar una rutina diaria en donde se pueden observar patrones repetitivos en las actividades realizadas en la semana.

Otro motivo por el cual se seleccionó un periodo de siete días, es porque se requiere experimentar y analizar los resultados para poder analizar la posibilidad de experimentar con la implementación de periodos mayores de monitoreo.

Se establece una rutina diaria “Normal” con duración a 7 días de lunes a domingo, organizando la rutina con 3 periodos establecidos seleccionados como: matutino, vespertino y nocturno. La rutina deberá ser realizada por una persona en las 4 habitaciones, cocina, comedor, sala y dormitorio.

Una persona realiza una rutina diaria “Normal” que está organizada como se muestra en la Tabla 5.32, (para consultar el significado del identificador de las actividades se puede consultar la **¡Error! No se encuentra el origen de la referencia.**).

Tabla 5.32. Entrenamiento del modelo: Rutina normal

Periodo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Matutino	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4
	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5
	B1,C3	B1,C1	B1,C3	B1,C1	B1,C3	C2	C2
Vespertino	A2,D2	A2,D2	A2,D2	A2,D2	A2,D2	A4	A4
	D2	D2	D2	D2	D2	C4	C4
	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1
Nocturno	B3,B1	B3,B1	B3,B1	D3	D3	C2	C2
	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1

Resultado de procesar un video en el entrenamiento

El modelo procesa un video de las actividades monitoreadas en una habitación para la fase de entrenamiento. La información que proporciona el modelo al terminar de procesar un video son los siguientes datos mostrados en la Tabla 5.33, mismos que son almacenados en una base de datos para ir generando un histórico de la rutina de una persona monitoreada.

Tabla 5.33. Entrenamiento: Datos resultado de procesar un video

Dato	Significado
Día	Corresponde al día de la rutina en el que se procesa el video
Periodo	Corresponde al periodo de la rutina en el que se procesa el video
Habitación	Corresponde a la habitación en la que se graba el video
Lista de actividades e identificador	Corresponde al resultado de reconocer la actividad realizada en el video
Etiquetas	Corresponde al identificador asignado a la actividad reconocida en el video
Conteo de etiquetas	Corresponde al número de etiquetas obtenidas resultado de reconocer la actividad realizada en el video
Secuencia de Actividades	Corresponde a la secuencia de actividades que fue reconocida en el video
Patrón	Corresponde a los caracteres asociados a la actividad reconocida en el video
Valor de la secuencia	Corresponde al valor del recorrido del grafo resultado del reconocimiento de actividades

Una persona entra a la habitación cocina, realiza sus actividades diarias y el modelo reconoce la siguiente secuencia de actividades: A1, A3, A5. Para la visualización de resultados, se divide en las siguientes Figura 5.45 y Figura 5.46.

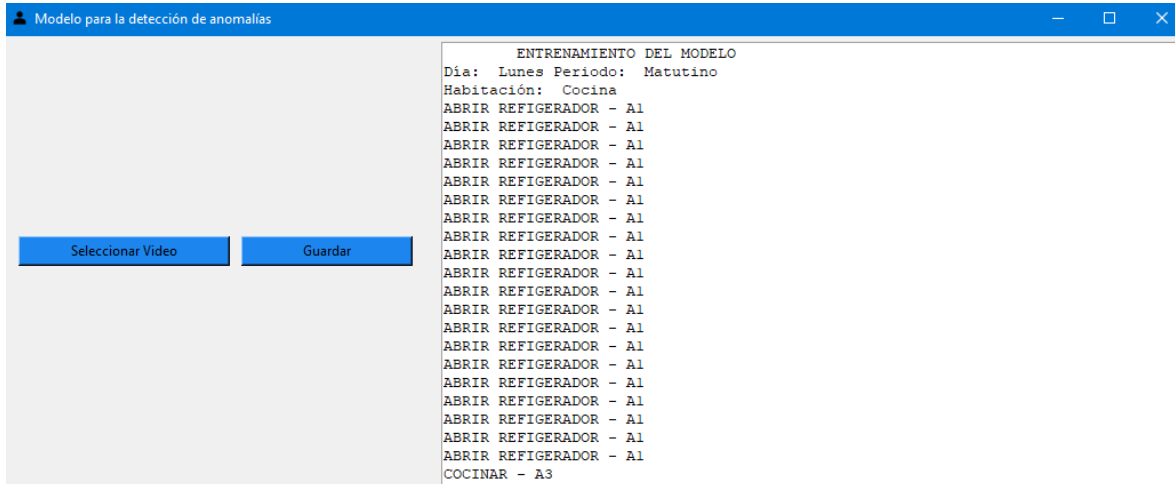


Figura 5.45. Resultado de procesar un video.

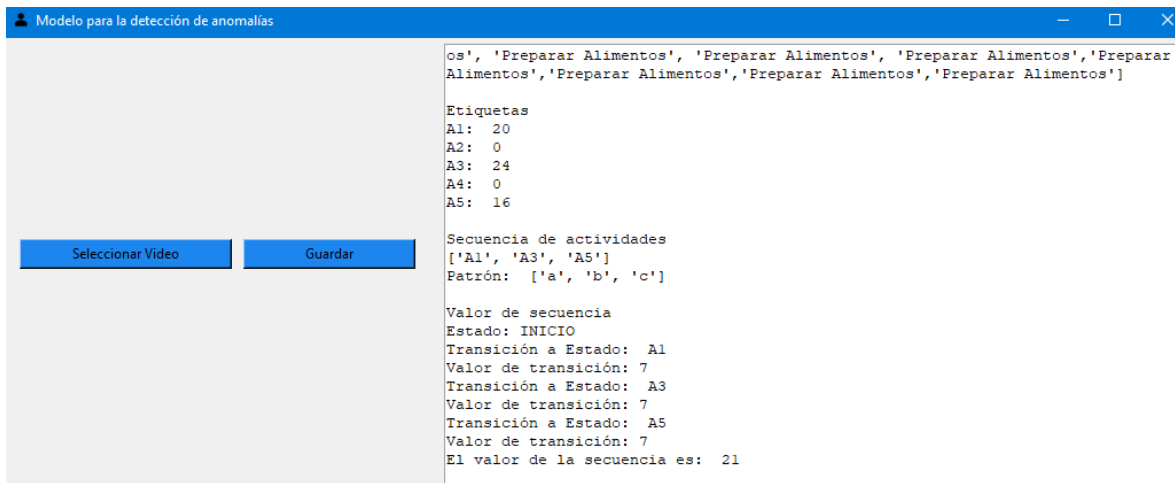


Figura 5.46. Resultado de procesar un video.

Visualizar imagen resultado

El sistema captura imágenes de las actividades reconocidas por el modelo en el video que fue procesado; posterior a la presentación de los datos resultado del video mencionados en la Tabla 5.33, se presentan las imágenes de las actividades reconocidas en la Figura 5.47.

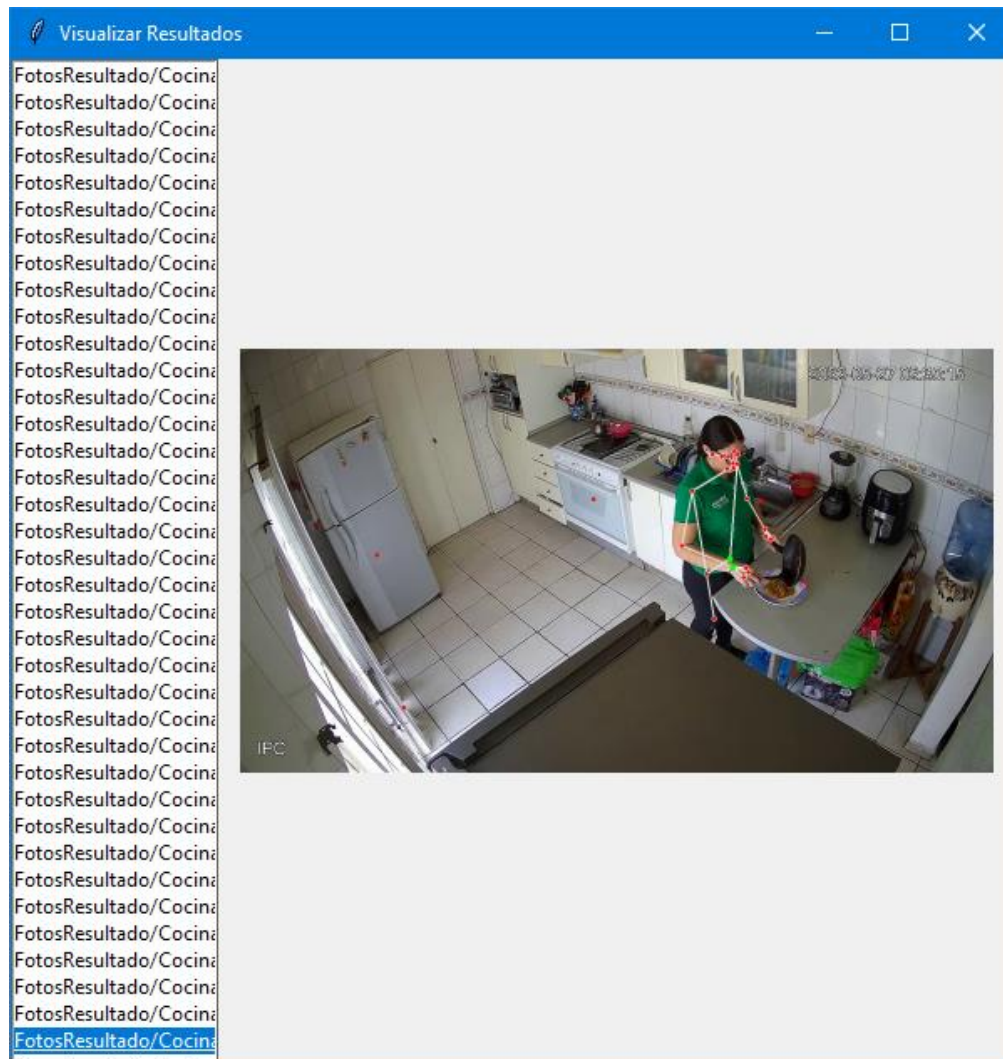


Figura 5.47. Visualizador de imágenes reconocidas por el modelo

5.11.2 Validación del modelo

Cuando el modelo se ha entrenado con los videos de una rutina normal ha terminado su proceso de entrenamiento, para después adquirir los datos de una rutina actual y realizar la fase de validación del modelo.

En la fase de validación, el modelo procesa un video y proporciona como resultado los datos descritos en la Tabla 5.34. Estos mismos datos son almacenados en una base de datos para generar un histórico de la rutina monitoreada de la persona.

Tabla 5.34. Validación: Datos resultado de procesar un video.

Dato	Significado
Día	Corresponde al día de la rutina en el que se procesa el video
Periodo	Corresponde al periodo de la rutina en el que se procesa el video
Habitación	Corresponde a la habitación en la que se graba el video
Lista de actividades e identificador	Corresponde al resultado de reconocer la actividad realizada en el video
Etiquetas	Corresponde al identificador asignado a la actividad reconocida en el video
Conteo de etiquetas	Corresponde al número de etiquetas obtenidas resultado de reconocer la actividad realizada en el video
Secuencia de Actividades	Corresponde a la secuencia de actividades que fue reconocida en el video
Patrón	Corresponde a los caracteres asociados a la actividad reconocida en el video
Valor de la secuencia	Corresponde al valor del recorrido del grafo resultado del reconocimiento de actividades
Validación por día	Indica si las actividades reconocidas en el video son válidas para el día en el que se procesa el video.
Validación por periodo	Indica si las actividades reconocidas en el video son válidas para el período en el que se procesa el video.
Clasificación del patrón	Indica si la cadena de caracteres asociados a las actividades reconocidas en el video, son aceptadas por la máquina de estado finito correspondiente.

Prueba 1:

Una persona realizó una secuencia de actividades dentro de una misma habitación que no fue aprendida por el modelo en la fase de entrenamiento, un ejemplo es C2, C3 que indica que la persona está viendo la TV y procede a leer el periódico, los resultados obtenidos se pueden consultar en la Figura 5.48 y en la Figura 5.49, resaltando que el patrón de entrada “no conoce una transición para el estado y la entrada” y reconoce el patrón como “anormal”.

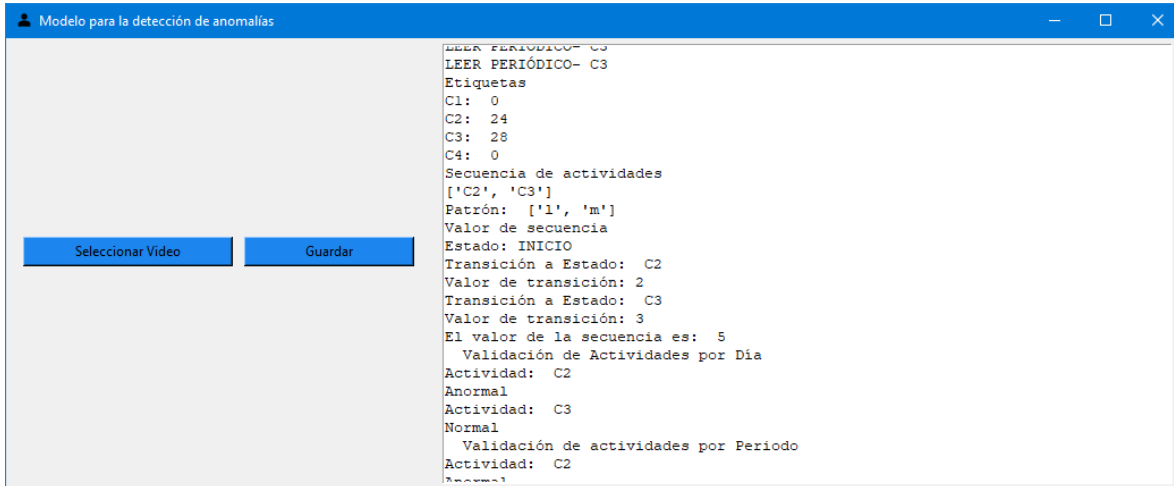


Figura 5.48. Resultado prueba 1.

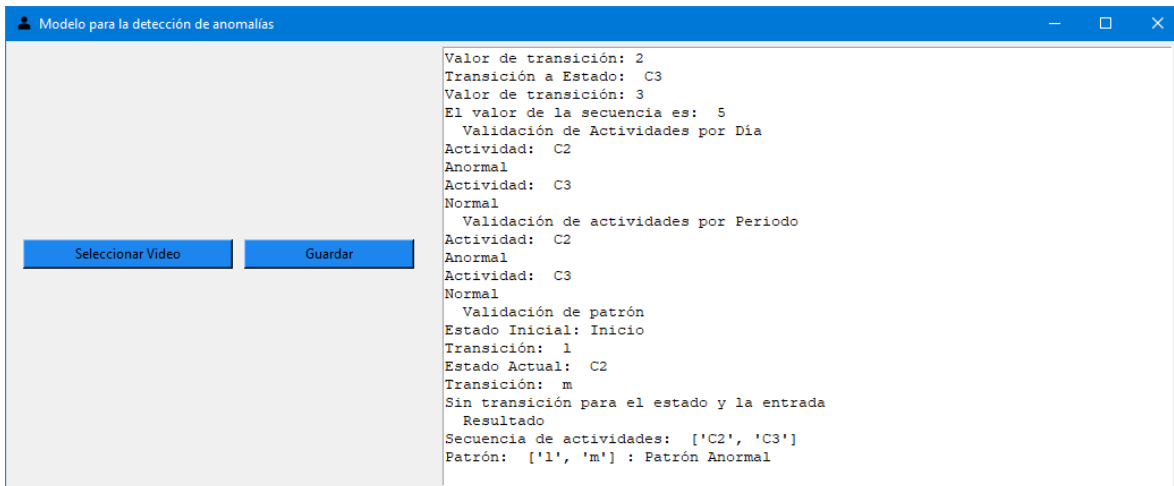


Figura 5.49. Resultado Prueba 1.

Prueba 2

Una persona realiza su rutina diaria normal, sin efectuar un cambio en el orden de sus actividades conforme a lo establecido en la Tabla 5.35 para validar que el modelo detecte que su comportamiento es “normal”, comparándolo con la primera rutina monitoreada y los datos de su histórico.

Rutina realizada

Tabla 5.35. Prueba 2: Rutina de actividades

Periodo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Matutino	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4
	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5
	B1,C3	B1,C1	B1,C3	B1,C1	B1,C3	C2	C2
Vespertino	A2,D2	A2,D2	A2,D2	A2,D2	A2,D2	A4	A4
	D2	D2	D2	D2	D2	C4	C4
	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1
Nocturno	B3,B1	B3,B1	B3,B1	D3	D3	C2	C2
	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1

Para visualizar los resultados de manera resumida, se mostrará el reconocimiento del patrón de comportamiento del día lunes en un archivo formato .txt mostrado en la Figura 5.50, para la descripción de cómo se realiza el reconocimiento de un patrón de comportamiento.

```

*AnálisisResultado.txt: Bloc de notas
Archivo Edición Formato Ver Ayuda
Día: Lunes
Patrón Actual: ['o', 's', 'a', 'b', 'c', 'f', 'm', 'd', 'p', 'g', 'h', 'i', 'j', 'n', 's']

Estado Inicial: Inicio
Transición: o
Estado Actual: D1
Transición: s
Estado Actual: D4
Transición: a
Estado Actual: A1
Transición: b
Estado Actual: A3
Transición: c
Estado Actual: A5
Transición: f
Estado Actual: B1
Transición: m
Estado Actual: C3
Transición: d
Estado Actual: A2
Transición: p
Estado Actual: D2
Transición: p
Estado Actual: D2
Transición: g
Estado Actual: B2
Transición: h
Estado Actual: B1
Transición: i
Estado Actual: B3
Transición: j
Estado Actual: B1
Transición: r
Estado Actual: D4
Transición: s
Estado Final: D1

Resultado
Patrón: Normal
Patrón válido en el día: Normal

Día: Martes
Patrón Actual: ['o', 's', 'a', 'b', 'c', 'f', 'k', 'd', 'p', 'g', 'h', 'i', 'j', 'n', 's']

```

Figura 5.50. Resultado de prueba 2 en formato txt

Se muestra en la interfaz gráfica del sistema el resultado final del reconocimiento de todos los días de la rutina (ver Figura 5.51). Se realiza un análisis de resultados en donde se indica

un conteo del total de registros (videos procesados por el modelo) que son clasificados como patrones normales, en esta prueba fueron 6 y un total de 0 patrones anormales. Además de indicar la clasificación del patrón de comportamiento para cada día de la semana.

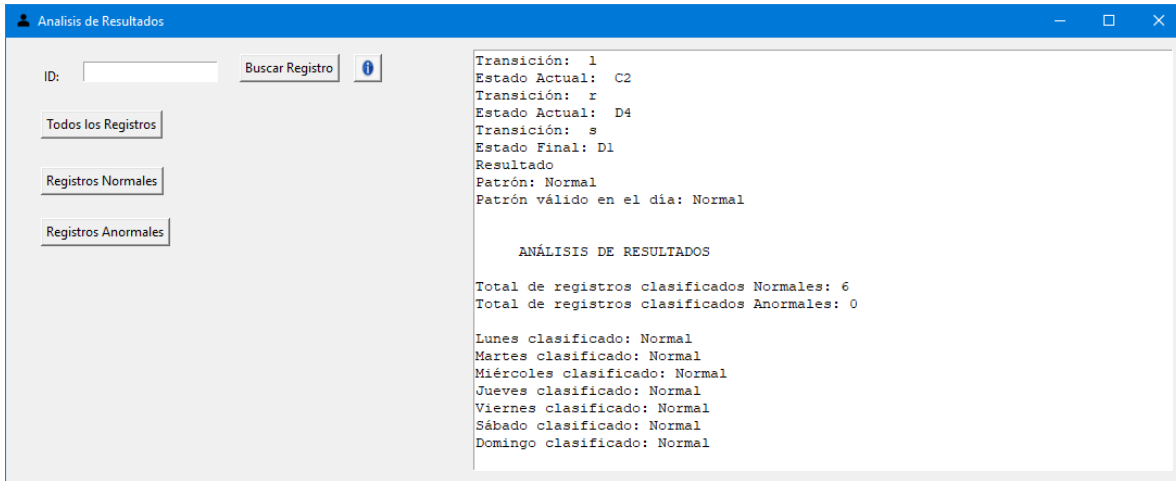


Figura 5.51. Resultado de prueba 2 en interfaz gráfica

Prueba 3

Una persona modifica su rutina diaria con la actividad D1 como se indica en la Tabla 5.36, que corresponde a la actividad “dormir” en el turno vespertino, en los días lunes, martes y miércoles. El modelo deberá ser capaz de reconocer esta desviación de su rutina diaria.

Tabla 5.36. Prueba 3: Rutina de actividades

Periodo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Matutino	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4
	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5
	B1,C3	B1,C1	B1,C3	B1,C1	B1,C3	C2	C2
Vespertino	A2,D2	A2,D2	A2,D2	A2,D2	A2,D2	A4	A4
	D1	D1	D1	D2	D2	C4	C4
	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1
Nocturno	B3,B1	B3,B1	B3,B1	D3	D3	C2	C2
	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1

Para visualizar los resultados de manera resumida, se muestra el reconocimiento del patrón de comportamiento del día lunes en un archivo formato .txt, (ver Figura 5.52).

```

AnalisisResultado3.txt: Bloc de notas
Archivo Edición Formato Ver Ayuda

Día: Lunes
Patrón Actual: ['o', 's', 'a', 'b', 'c', 'f', 'm', 'd', 'p', 'o', 'g', 'h', 'i', 'j', 'n', 's']
Estado Inicial: Inicio
Transición: o
Estado Actual: D1
Transición: s
Estado Actual: D4
Transición: a
Estado Actual: A1
Transición: b
Estado Actual: A3
Transición: c
Estado Actual: A5
Transición: f
Estado Actual: B1
Transición: m
Estado Actual: C3
Transición: d
Estado Actual: A2
Transición: p
Estado Actual: D2
Transición: o

Sin transición para el estado y la entrada
Realizó una secuencia de actividades no aprendida
Resultado
Patrón Anormal

Anormalidad al realizar la actividad: Dormir

La actividad se realiza con normalidad los días
Lunes
Martes
Miércoles
Jueves
Viernes
Sabado
Domingo

La actividad se realiza con normalidad en el período
Matutino
Nocturno
Ln 4, Col 1 100% Windows (CRLF) UTF-8

```

Figura 5.52. Resultado de prueba 3 formato txt

Se muestra en la interfaz gráfica del sistema el resultado final del reconocimiento de todos los días de la rutina (ver Figura 5.53). Como resultado se muestra un conteo total de 4 registros y un total de 0 patrones anormales. Además de indicar la clasificación del patrón de comportamiento para cada día de la semana, donde se indica el reconocimiento de un patrón de comportamiento anormal en el día lunes, martes y miércoles.

```

Análisis de Resultados

ID:  Buscar Registro ⓘ

Todos los Registros
Registros Normales
Registros Anormales

Estado Actual: B1
Transición: l
Estado Actual: C2
Transición: r
Estado Actual: D4
Transición: s
Estado Final: D1
Resultado
Patrón: Normal
Patrón válido en el día: Normal

ANÁLISIS DE RESULTADOS
Total de registros clasificados Normales: 4
Total de registros clasificados Anormales: 3
Lunes clasificado: Anormal
Martes clasificado: Anormal
Miércoles clasificado: Anormal
Jueves clasificado: Normal
Viernes clasificado: Normal
Sábado clasificado: Normal
Domingo clasificado: Normal

```

Figura 5.53. Resultado de prueba 3 en interfaz gráfica

Prueba 4

Una persona realiza todas las actividades programadas en el día domingo en su rutina “normal”, después en su rutina actual, realizándolas en el día jueves como en la Tabla 5.37. El sistema deberá ser capaz de detectar este cambio en la rutina de la persona.

Tabla 5.37. Prueba 4: Rutina de actividades

Periodo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Matutino	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4
	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5
	B1,C3	B1,C1	B1,C3	C2	B1,C3	C2	C2
Vespertino	A2,D2	A2,D2	A2,D2	A4	A2,D2	A4	A4
	D4	D4	D4	C4	D2	C4	C4
	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1
Nocturno	B3,B1	B3,B1	B3,B1	C2	D3	C2	C2
	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1

Se muestra el reconocimiento del patrón de comportamiento del día Jueves en formato .txt (ver Figura 5.54), debido a la extensión de los resultados, sin embargo, se resalta que el resultado del reconocimiento también es proporcionado en la interfaz gráfica del sistema.

Como se puede observar en el contenido del archivo, el sistema va ejecutando el patrón de comportamiento específicamente del día de interés en esta prueba, como resultado obtiene que se trata de un patrón de comportamiento normal, pero en la validación del patrón acorde al día de la semana, el modelo detecta que no se realizaron las actividades que suele hacer en el día jueves, por lo tanto, realiza la clasificación del patrón anormal al día en el que se realizó.

```

*AnálisisResultados_4.txt: Bloc de notas
Archivo Edición Formato Ver Ayuda
Día: Jueves
Patrón Actual: ['o', 's', 'a', 'b', 'c', 'l', 'e', 'n', 'g', 'h', 'l', 'e', 's']

Estado Inicial: Inicio
Transición: o
Estado Actual: D1
Transición: s
Estado Actual: D4
Transición: a
Estado Actual: A1
Transición: b
Estado Actual: A3
Transición: c
Estado Actual: A5
Transición: l
Estado Actual: C2
Transición: e
Estado Actual: A4
Transición: n
Estado Actual: C4
Transición: g
Estado Actual: B2
Transición: h
Estado Actual: B1
Transición: l
Estado Actual: C2
Transición: r
Estado Actual: D4
Transición: s
Estado Final: D1

Resultado
Patrón: Normal

Patrón válido en el día: Anormal
|
Patrón Normal habitualmente en día:
Sábado
Domingo

```

Figura 5.54. Resultado de prueba 4 en formato txt

Prueba 5

Una persona realiza la secuencia de actividades programadas para el día sábado y domingo en su rutina “normal”, ahora en su rutina actual realizada en los días lunes y martes y viceversa en la Tabla 5.38.

Tabla 5.38. Prueba 5: Rutina de actividades

Periodo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Matutino	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4	D1, D4
	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5	A1,A3,A5
	C2	C2	B1,C3	B1,C1	B1,C3	B1,C3	B1,C1
Vespertino	A4	A4	A2,D2	A2,D2	A2,D2	A2,D2	A2,D2
	C4	C4	D4	D2	D2	D2	D2
	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1	B2,B1
Nocturno	C2	C2	B3,B1	D3	D3	B3,B1	B3,B1
	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1	D4,D1

Se muestra el reconocimiento del patrón de comportamiento del día jueves en el archivo formato .txt mostrado en la Figura 5.55, debido a la extensión de los resultados, sin embargo, se resalta que el resultado del reconocimiento también es proporcionado en la interfaz gráfica del sistema.

Los resultados indican que se encontró un total de 3 registros con patrones normales y un total de 4 registros clasificados como anormales. Además, se muestra en que día de la semana se detectó un patrón anormal y como resultado se trató de los días lunes, martes, sábado y domingo.

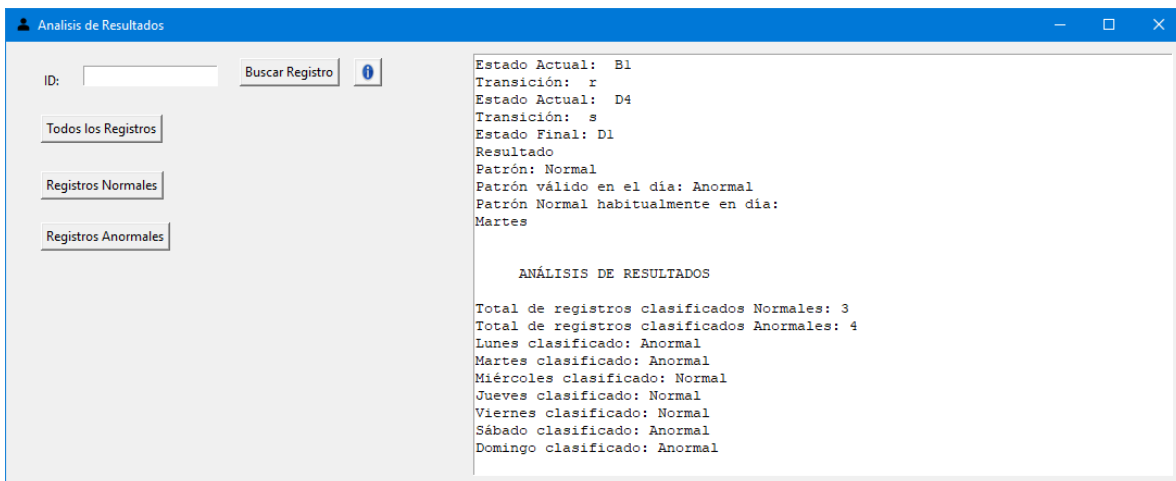


Figura 5.55. Resultado de prueba 5 en interfaz gráfica.

5.12 Discusión

De acuerdo con los resultados obtenidos se pueden resumir lo siguiente de cada una de las fases que conforman el procesamiento de la información del modelo.

Se comienza con la etapa de localización de la persona, en donde los resultados permiten deducir qué el método más óptimo para la localización de una persona es la implementación de la aplicación móvil puesto que obtuvo mejores resultados.

En la etapa de la estimación de la pose humana resultó más favorable haber utilizado el método de *Mediapipe pose* donde el tiempo de procesamiento permite implementar la estimación de la pose en tiempo real y con un menor coste de estimación. 7

En la etapa de reconocimiento de objetos con la librería YOLO V3 se obtiene un alto valor de confianza de predicción de los objetos detectados en las habitaciones, por ejemplo (refrigerador 99%, estufa 94%, sofá 95%, etc.). En los casos donde la confianza de predicción es baja, se realizó un etiquetado manual de la zona de interés.

En la etapa de recolección de datos del sensor se realizó un experimento adicional utilizando un modelo de red neuronal profunda (DNN) con resultados del banco de datos WISDM del 74% y el conjunto de datos propio con 65% que indican que la manera en la que se está monitoreando a una persona a través de los sensores es con la frecuencia de muestreo y posición adecuada.

En la etapa de reconocimiento de actividad se obtuvo un mejor porcentaje de *accuracy* en las actividades cocinar, preparar alimentos y usar tocador, en el caso de las actividades salir de casa, usar microondas y leer periódico existe un desafío por parte de la metodología de *mediapipe pose*, sin embargo, la mayoría de las actividades presentaron un porcentaje de exactitud confiable.

En la etapa de modelado de la rutina se generaron en total seis grafos dirigidos cada uno con estados y transiciones derivados del procesamiento de la rutina de una persona, en caso de querer modificar las actividades de la rutina de una persona, el modelo tiene la posibilidad de adaptarse a nuevas modificaciones, como la posibilidad de modificar actividades y modificar número de habitaciones.

En la etapa de reconocimiento de anomalías se implementaron máquinas de estado finito que reciben como entrada un patrón de comportamiento, las máquinas evaluaron correctamente el patrón de entrada y fueron capaces de detectar desviaciones en el comportamiento de una persona.

Capítulo 6

Análisis de resultados, conclusiones y trabajo futuro

En el capítulo final de este documento se presenta un análisis de los resultados obtenidos, la conclusión final, los aportes de la investigación y las áreas de oportunidad para desarrollar un trabajo a futuro de esta investigación.

6.1 Análisis de resultados

En esta sección de análisis de resultados se describe el significado de los seis eventos reconocidos por el modelo, como es el caso de reconocer una anomalía en el día y periodo de la rutina de una persona, y el resto de los eventos clasificados por el modelo, descritos en la Tabla 6.1.

Tabla 6.1. Definición de eventos clasificados por el modelo

Etiqueta	Definición
Anormalidad en el día	Significa que se ha encontrado al menos una actividad que no es realizada habitualmente en el día registrado
Anormalidad en el periodo	Significa que se ha encontrado al menos una actividad que no es realizada habitualmente en el periodo registrado
Patrón anormal	Significa cadena de caracteres no aceptada, se traduce a secuencia de actividades que no son una rutina aprendida por el modelo
Patrón normal	Significa cadena de caracteres si aceptada, se traduce a secuencia de actividades aprendida por el modelo

Tabla 6.1. Definición de eventos clasificados por el modelo (Continuación)

Etiqueta	Definición
Patrón de comportamiento anormal al día	Significa que ha realizado una secuencia de actividades en todo el día conocida por el modelo, pero esta secuencia de actividades realizada no corresponde al día registrado.
Patrón inhabitual	Significa cadena de caracteres no aceptada, se traduce a relación entre secuencia de actividades nunca aprendida antes por el modelo

De acuerdo con los resultados obtenidos de las pruebas realizadas, se puede deducir que el modelo propuesto cumple con los objetivos planteados para esta investigación. El modelo logró el reconocimiento de objetos en la escena, implementó la estimación de la postura humana y después reconoció 16 actividades de la vida diaria, se realizó el monitoreo de una rutina durante 7 días y se monitoreó la movilidad de una persona al usar el sensor giroscopio y acelerómetro de un dispositivo móvil inteligente. Las actividades que mostraron mejor exactitud de reconocimiento fueron: cocinar con 95% de exactitud, preparar alimentos 95%, usar tocador 95%. Las actividades que presentaron mayor complejidad de reconocimiento fueron: salir de casa con 85%, usar microondas 89% y leer periódico 89%, puesto que se identificó una dificultad en la metodología implementada. Finalmente, el modelo logra el reconocimiento de un comportamiento “anormal” en la rutina de una persona, además de informar el tipo de anomalía presente en la rutina como es el caso de identificar una anomalía realizada en un día de la rutina, una anomalía presente en el periodo de la rutina o el reconocimiento de un patrón inhabitual.

6.2 Conclusión

En el desarrollo de este modelo, para la detección de anomalías en el comportamiento de una persona, se ha observado que es posible modelar un comportamiento “normal” y en base con la experimentación realizada, los resultados obtenidos permiten, el enriquecimiento del modelo. Se ha logrado observar, métricas que deben ser consideradas para hacer la detección, de un evento “anormal” en el comportamiento de una persona. Con el avance de la tecnología, la propuesta de este modelo para la detección de anomalías, podría aplicarse a una amplia variedad de situaciones en la vida real, como, por ejemplo, dentro de un entorno escolar, monitoreo de adultos mayores o monitoreo de personas que requieran una vida independiente dentro del hogar.

6.3 Aportaciones

Este trabajo de investigación tiene como mayor aporte la metodología que es un proceso para la detección de anomalías en el comportamiento, en este trabajo en particular de una persona.

En el proceso se involucran los siguientes pasos:

1. Adquirir videos de cuatro cámaras RGB.
2. Realizar la detección de objetos en video.
3. Realizar la localización de una persona y grabar video.
4. Adquirir datos del sensor acelerómetro y sensor giroscopio.
5. Realizar la estimación de la pose humana.
6. Realizar el cálculo de medida de proximidad.
7. Implementar grafos dirigidos y máquinas de estado finito.
8. Reconocimiento de actividad.
9. Modelar rutina.
10. Reconocimiento de anomalía.

Adicionalmente, cabe resaltar que se publicó un artículo que fue producto derivado de esta investigación (consultar **Anexo B**).

6.4 Trabajos futuros

Como continuación de este trabajo de tesis, existen diversas líneas de investigación que quedan abiertas y en las que es posible continuar trabajando. A continuación, se presentan algunos trabajos futuros que pueden desarrollarse como resultado de esta investigación que no han podido ser tratados con la suficiente profundidad. Entre los posibles trabajos futuros se destacan:

- Extender periodos de monitoreo de la rutina de una persona, que permitan obtener mayores datos sobre las actividades de una persona para modelar de mejor manera un comportamiento.
- Incluir el calendario laboral oficial de México, que contiene los días festivos y fechas importantes de México, que permita considerar aquellos días inhábiles, para considerar esta información al momento de reportar una “anomalía” en el comportamiento de una persona.
- El modelo propuesto sea capaz de volver a aprender una rutina “normal”, esto debido a que la persona monitoreada se encuentre en una situación en donde su rutina diaria ha cambiado y esto no sea informado como una anomalía en su comportamiento actual

Referencias

- Abayomi, A., Olugbara, O. O., & Heukelman, D. (2018). An Architecture Utilizing Human Emotions and Activities Recognition for Remote Monitoring. *2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*, (págs. 1–4).
- Aggarwal, J. K., & Ryoo, M. S. (4 de 2011). Human Activity Analysis: A Review. *ACM Comput. Surv.*, *43*. doi:10.1145/1922649.1922653
- Alshurafa, N., Eastwood, J., Nyamathi, S., Liu, J. J., Xu, W., Ghasemzadeh, H., . . . Sarrafzadeh, M. (2015). Improving Compliance in Remote Healthcare Systems Through Smartphone Battery Optimization. *IEEE Journal of Biomedical and Health Informatics*, *19*, 57-63. doi:10.1109/JBHI.2014.2329712
- Amrutha, K., Prabu, P., & Paulose, J. (2021). Human Body Pose Estimation and Applications. *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, (págs. 1-6). doi:10.1109/i-PACT52855.2021.9696513
- Avci, A., Bosch, S., Marin-Perianu, M., Marin-Perianu, R., & Havinga, P. (2010). Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey. In *23th International Conference on Architecture of Computing Systems 2010* (pp. 1-10).
- Babiker, M., Khalifa, O. O., Htike, K. K., Hassan, A., & Zaharadeen, M. (2017). Automated daily human activity recognition for video surveillance using neural network. *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, (págs. 1-5). doi:10.1109/ICSIMA.2017.8312024
- Barrett, S. F. *Arduino Microcontroller Processing for Everyone!: Third Edition*, 2013
- Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). BlazePose: On-device Real-time Body Pose tracking. *BlazePose: On-device Real-time Body Pose tracking*. arXiv. doi:10.48550/ARXIV.2006.10204
- Bobick, A. F., & Davis, J. W. (2001). The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*, 257-267. doi:10.1109/34.910878
- Bruno, B., Mastrogiovanni, F., Sgorbissa, A., Vernazza, T., & Zaccaria, R. (2013). Analysis of human behavior recognition algorithms based on acceleration data. *2013 IEEE International Conference on Robotics and Automation*, (págs. 1602-1607). doi:10.1109/ICRA.2013.6630784

- Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43, 172-186. doi:10.1109/TPAMI.2019.2929257
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. *CoRR*, abs/2005.12872. Obtenido de <https://arxiv.org/abs/2005.12872>
- Chakraborty, S. (2003). Formal Languages and Automata Theory- Regular Expressions and Finite Automata- Contents. *Formal Languages and Automata Theory- Regular Expressions and Finite Automata- Contents*.
- Chen, C., Jafari, R., & Kehtarnavaz, N. (2015). UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. *2015 IEEE International Conference on Image Processing (ICIP)*, (págs. 168-172). doi:10.1109/ICIP.2015.7350781
- Chen, Y.-P., Yang, J.-Y., Liou, S.-N., Lee, G.-Y., & Wang, J.-S. (2008). Online classifier construction algorithm for human activity detection using a tri-axial accelerometer. *Applied Mathematics and Computation*, 205, 849-860. doi:<https://doi.org/10.1016/j.amc.2008.05.099>
- Chunduru, V., Roy, M., Dasari Romit N., S., & Chittawadigi, R. G. (2021). Hand Tracking in 3D Space using MediaPipe and PnP Method for Intuitive Control of Virtual Globe. *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)*, (págs. 1-6). doi:10.1109/R10-HTC53172.2021.9641587
- Du, X., Lin, T.-Y., Jin, P., Ghiasi, G., Tan, M., Cui, Y., . . . Song, X. (2020). SpineNet: Learning Scale-Permuted Backbone for Recognition and Localization. *CVPR*.
- D. U. R. Castellanos, “Seguimiento y caracterización automática de la trayectoria de una persona, en un entorno multicámaras de videovigilancia,” *Master’s thesis, CENIDET, Cuernavaca, Morelos, Enero 2019*.
- Duong, T. V., Bui, H. H., Phung, D. Q., & Venkatesh, S. (2005). Activity recognition and abnormality detection with the switching hidden semi-Markov model. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 1, págs. 838-845 vol. 1. doi:10.1109/CVPR.2005.61
- Ganesh, P., Idgahi, R. E., Venkatesh, C. B., Babu, A. R., & Kyrarini, M. (2020). Personalized System for Human Gym Activity Recognition Using an RGB Camera. *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. New York, NY, USA: Association for Computing Machinery. doi:10.1145/3389189.3392611
- Garcia-Gonzalez, D., Rivero, D., Fernandez-Blanco, E., & Luaces, M. R. (2020). A Public Domain Dataset for Real-Life Human Activity Recognition Using Smartphone Sensors. *Sensors*, 20. doi:10.3390/s20082200

- Gaglio, G. L. Re, & M. Morana (2015). Human Activity Recognition Process Using 3-D Posture Data. *IEEE Transactions on Human-Machine Systems*, 45(5), 586-597.
- Helaoui, R., Niepert, M., & Stuckenschmidt, H. (2011). Recognizing interleaved and concurrent activities: A statistical-relational approach. *2011 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, (págs. 1-9). doi:10.1109/PERCOM.2011.5767586
- Hossen, M. K., & Tuli, S. H. (2016). A surveillance system based on motion detection and motion estimation using optical flow. *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, (págs. 646-651). doi:10.1109/ICIEV.2016.7760081
- Hu, X., Li, Z., Dai, R., Cui, Y., Zhou, Z., An, B., . . . Chu, D. (2020). Coarse-to-Fine Activity Annotation and Recognition Algorithm for Solitary Older Adults. *IEEE Access*, 8, 4051-4064. doi:10.1109/ACCESS.2019.2962843
- Jennifer R. Kwapisz, Gary M. Weiss and Samuel A. Moore (2010). Activity Recognition using Cell Phone Accelerometers, Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data (at KDD-10), Washington DC
- Joon-Ho Lim and Hyunchul Jang and Jaewon Jang and S. Park. (2008). Daily activity recognition system for the elderly using pressure sensors. *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (págs. 5188-5191). doi:10.1109/IEMBS.2008.4650383
- Krause, E. F. (1975). Taxicab geometry: An adventure in Non-Euclidean geometry. New York: Dover Publications.
- Lara, Ó., Perez, A., Labrador, M., & Posada, J. (10 de 2012). Centinela: A human activity recognition system based on acceleration and vital sign data. *Pervasive and Mobile Computing*, 8, 717. doi:10.1016/j.pmcj.2011.06.004
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. *Focal Loss for Dense Object Detection*. arXiv. doi:10.48550/ARXIV.1708.02002
- Mahendru, M., & Dubey, S. K. (2021). Real Time Object Detection with Audio Feedback using Yolo vs. Yolo_v3. *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, (págs. 734-740). doi:10.1109/Confluence51648.2021.9377064
- Mahjoub, A. B., & Atri, M. (2016). Human action recognition using RGB data. *2016 11th International Design Test Symposium (IDT)*, (págs. 83-87). doi:10.1109/IDT.2016.7843019
- Mahmud, T., Akash, S. S., Fattah, S. A., Zhu, W. P., & Ahmad, M. O. (2020). Human Activity Recognition From Multi-modal Wearable Sensor Data Using Deep Multi-stage LSTM Architecture Based on Temporal Feature Aggregation. *2020 IEEE 63rd*

- International Midwest Symposium on Circuits and Systems (MWSCAS)*, (págs. 249-252). doi:10.1109/MWSCAS48704.2020.9184666
- Maurer, U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006). Activity recognition and monitoring using multiple sensors on different body positions. *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)*, (págs. 4 pp.-116). doi:10.1109/BSN.2006.6
- Mettel, M., Alekseew, M., Stocklów, C., & Braun, A. (2 de 2019). Designing and evaluating safety services using depth cameras. *Journal of Ambient Intelligence and Humanized Computing*, *10*. doi:10.1007/s12652-018-0898-9
- M. L. D. García, “Caracterización automática de una muestra de acciones en video,” *Master’s thesis, CENIDET, Cuernavaca, Morelos, Diciembre 2019*.
- Novac, O. C., Novac, M., Gordan, C., Berczes, T., & Bujdosó, G. (2017). Comparative study of Google Android, Apple iOS and Microsoft Windows Phone mobile operating systems. *2017 14th International Conference on Engineering of Modern Electric Systems (EMES)*, (págs. 154-159). doi:10.1109/EMES.2017.7980403
- Oikonomopoulos, A., Patras, I., & Pantic, M. (2006). Spatiotemporal salient points for visual recognition of human actions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *36*, 710-719. doi:10.1109/TSMCB.2005.861864
- Pavel, M., Jimison, H. B., Korhonen, I., Gordon, C. M., & Saranummi, N. (2015). Behavioral Informatics and Computational Modeling in Support of Proactive Health Management and Care. *IEEE Transactions on Biomedical Engineering*, *62*, 2763-2775. doi:10.1109/TBME.2015.2484286
- Redmon, J. & Farhadi, A. YOLOv3: An Incremental Improvement arXiv, 2018
- Riboni, D., & Bettini, C. (3 de 2011). COSAR: Hybrid Reasoning for Context-aware Activity Recognition. *Personal and Ubiquitous Computing*, *15*, 271-289. doi:10.1007/s00779-010-0331-7
- R. A. Juárez, “Caracterización visual de movimientos sospechosos de personas en estacionamientos,” *Master’s thesis, CENIDET, Cuernavaca, Morelos, Marzo 2013*.
- Skubic, M., Guevara, R. D., & Rantz, M. (2015). Automated Health Alerts Using In-Home Sensor Data for Embedded Health Assessment. *IEEE Journal of Translational Engineering in Health and Medicine*, *3*, 1-11. doi:10.1109/JTEHM.2015.2421499
- Sreenath, S., Daniels, D. I., Ganesh, A. S., Kuruganti, Y. S., & Chittawadigi, R. G. (2021). Monocular Tracking of Human Hand on a Smart Phone Camera using MediaPipe and its Application in Robotics. *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)*, (págs. 1-6). doi:10.1109/R10-HTC53172.2021.9641542

- Thummala, J., & Pumrin, S. (2020). Fall Detection using Motion History Image and Shape Deformation. *2020 8th International Electrical Engineering Congress (iEECON)*, (págs. 1-4). doi:10.1109/iEECON48109.2020.229491
- Ul-Haq, M. E., Javed, A., Azam, M. A., Malik, H. M., Irtaza, A., Lee, I. H., & Mahmood, M. T. (2019). Robust Human Activity Recognition Using Multimodal Feature-Level Fusion. *IEEE Access*, 7, 60736-60751. doi:10.1109/ACCESS.2019.2913393
- Van Kasteren, T., Noulas, A., Englebienne, G., & Kröse, B. (2008). Accurate Activity Recognition in a Home Setting. *Proceedings of the 10th International Conference on Ubiquitous Computing* (págs. 1–9). New York, NY, USA: Association for Computing Machinery. doi:10.1145/1409635.1409637
- Vanus, J. a. (2017). Monitoring of the daily living activities in smart home care. *Human-centric Computing and Information Sciences*.
- Vishwakarma, S., & Agrawal, A. (2013). A survey on activity recognition and behavior understanding in video surveillance. *The Visual Computer*, 29, 983–1009.
- Wasserman, S., Faust, K., & Molina, J. L. (2013). *Análisis de redes sociales. Métodos y aplicaciones*. Centro de Investigaciones Sociológicas. Obtenido de https://books.google.com.mx/books?id=Da_{T}{A}{g}AAQBAJ
- Weiss, G. M., Yoneda, K., & Hayajneh, T. (2019). Smartphone and Smartwatch-Based Biometrics Using Activities of Daily Living. *IEEE Access*, 7, 133190-133202. doi:10.1109/ACCESS.2019.2940729
- Xiang, T., & Gong, S. (2005). Video behaviour profiling and abnormality detection without manual labelling. *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 2*, págs. 1238-1245 Vol. 2. doi:10.1109/ICCV.2005.248
- Yang, A., Jafari, R., Sastry, S., & Bajcsy, R. (1 de 2009). Distributed recognition of human actions using wearable motion sensor networks. *JAISE*, 1, 103-115. doi:10.3233/AIS-2009-0016
- Yin, J., Yang, Q., & Pan, J. J. (2008). Sensor-Based Abnormal Human-Activity Detection. *IEEE Transactions on Knowledge and Data Engineering*, 20, 1082-1090. doi:10.1109/TKDE.2007.1042
- Yuan, C., & Xu, W. (2011). Multi-object events recognition from video sequences using extended finite state machine. *2011 4th International Congress on Image and Signal Processing*, 1, págs. 202-205. doi:10.1109/CISP.2011.6099975
- Zhen-Yu He and Lian-Wen Jin. (2008). Activity recognition from acceleration data using AR model representation and SVM. *2008 International Conference on Machine Learning and Cybernetics*, 4, págs. 2245-2250. doi:10.1109/ICMLC.2008.4620779
- Zhu, C., & Sheng, W. (2009). Multi-sensor fusion for human daily activity recognition in robot-assisted living. *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, (págs. 303-304). doi:10.1109/ROBOT.2009.5152756

Anexo A. Open Pose

Se presenta las imágenes resultado de implementar la estimación de la pose humana con la librería Open Pose Tabla A.1.

Tabla A.1. Resultados de Open Pose









Actividad	Imagen
Cocinar	 A high-angle camera shot of a kitchen. A person is standing at the sink area, facing right. A multi-colored skeleton overlay is visible on the person's body. The kitchen has white cabinets, a sink, and a refrigerator. A timestamp in the top right corner reads '2022-05-27 12:18:33'. The text 'IPC' is visible in the bottom left corner.
Abrir refrigerador	 A high-angle camera shot of a kitchen, similar to the first image. The person is now standing in front of the refrigerator, which is open. The skeleton overlay is visible. A timestamp in the top right corner reads '2022-05-27 12:18:33'. The text 'IPC' is visible in the bottom left corner.
Comer	 A high-angle camera shot of a dining room. A person is sitting at a dark wooden dining table, eating. A skeleton overlay is visible on the person. The room has a window with blinds and a television in the background. A timestamp in the top right corner reads '2022-05-28 02:10:22'. The text 'IPC' is visible in the bottom left corner.
Tomar fruta	 A high-angle camera shot of a dining room, similar to the third image. The person is sitting at the table, reaching towards a bowl of fruit. A skeleton overlay is visible. A timestamp in the top right corner reads '2022-05-28 02:09:56'. The text 'IPC' is visible in the bottom left corner.

Tabla A.1. Resultados de Open Pose (Continuación)

Actividad	Imagen
Ver TV	
Hablar por teléfono	
Dormir	
Ejercitarse	
Trabajar en PC	

Anexo B. Artículo

Modelo para la detección de anomalías en el comportamiento de una persona

Gloria Vanessa Urquiza Flores, Raúl Pinto Elías *

*Departamento de Ciencias Computacionales
Tecnológico Nacional de México/CENIDET
{m20ce089, raul.pe}@cenidet.tecnm.mx*

Resumen: El reconocimiento de actividad humana HAR (Human Activity Recognition) es un área de investigación activa, el interés por desarrollar esta área se debe a su amplia variedad de aplicaciones en el mundo real como, en la videovigilancia, seguridad, deporte, atención médica, vida asistida y monitoreo continuo en el hogar. El HAR es un tema relacionado con el análisis del comportamiento humano con gran trayectoria de investigación, mayormente enfocado al comportamiento individual de un ser humano, monitoreando el comportamiento de una persona. En el presente artículo se propone el desarrollo de un modelo multimodal utilizando cámaras RGB, acelerómetro y giroscopio. Mediante la observación de lo que suele hacer una persona durante una rutina durante un periodo establecido, el modelo podrá utilizar su conocimiento aprendido para monitorear la rutina actual de la persona y detectar si existe alguna desviación de los patrones de actividad normales.

Palabras clave: Actividad humana, patrón de comportamiento, patrón anómalo.

1. INTRODUCCIÓN

El reconocimiento automático de actividades físicas es conocido como Reconocimiento de Actividad Humana HAR, su objetivo es el estudio y modelado de los movimientos que provienen del cuerpo humano, al realizar una actividad (J. Thummala, 2020). El HAR está desarrollado para aplicaciones como, la atención médica (M. Skubic, 2015), vida asistida (Mettel, M., 2019), vigilancia (M. K. Hossen, 2016), seguridad, deporte, entre otras.

Se utilizan métodos que implementan sensores para monitorear a los seres humanos y sus actividades diarias. Aunque cada sensor tiene ventajas y desventajas significativas, usar un solo tipo de sensor para monitorear actividades, ha demostrado ser insuficiente para caracterizar una amplia variedad de actividades físicas humanas.

El reconocimiento de actividades humanas está relacionado con el análisis de comportamiento humano, en la gran mayoría de estos sistemas, se enfocan en monitorear el comportamiento individual de una persona. Se realiza el procesamiento de secuencias de imágenes o videos en entornos controlados con el objetivo de obtener un modelo de comportamiento.

En el área de HAR existe un enfoque en el desarrollo de tecnología, que permita la evaluación continua de los cambios basados en el comportamiento y los patrones de actividad de un individuo, con el objetivo de determinar la existencia de cambios en el diario vivir de un ser humano.

La principal aportación de este artículo, es el desarrollo de un sistema multimodal utilizando los sensores cámaras RGB, acelerómetro y giroscopio. El escenario para el sistema son cuatro habitaciones controladas de un hogar y la caracterización de 16 actividades de la vida diaria para realizar el reconocimiento de patrones anómalos en el comportamiento de una sola persona.

El resto del documento se encuentra organizado de la siguiente forma: en la sección dos se presenta el estado del arte, en la sección tres se analizan los fundamentos teóricos de la investigación, en la sección cuatro se describe el modelo propuesto, en la sección cinco se detalla la experimentación y se discuten los resultados obtenidos. Finalmente, se presenta la sección de conclusiones.