

TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Culiacán



Opción:
TITULACIÓN INTEGRAL
TESIS PROFESIONAL

“Extracción de siluetas del caminar de una persona”

ALUMNO:
Oscar Oswaldo Meza León

CARRERA:
Ingeniería en Sistemas Computacionales

No. DE CONTROL:
17171427

Director:
MCC. Gloria Ekaterine Peralta Peñuñuri

Codirector:
Dr. Héctor Rodríguez Rangel

Culiacán, Sinaloa. Octubre, 2022.



Instituto Tecnológico de Culiacán

DIVISION DE ESTUDIOS PROFESIONALES
ASUNTO: Autorización de Impresión

Culiacán, Sinaloa, 10/11/2022 19:32

C. OSCAR OSWALDO MEZA LEÓN
PASANTE(S) DE LA CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES
P R E S E N T E.

De acuerdo con fallo emitido por la Comisión Revisora Integrada por:

	Cédula Profesional
Presidente: Maestría en Ciencias en Ciencias de la Computación Gloria Ekaterine Peralta Peñuñuri	6038275
Secretario: Doctorado en Ciencias en Ingeniería Eléctrica Opción en Sistemas Computacionales Héctor Rodríguez Rangel	9972835
Vocal: Licenciado en Informática Nora Esmeralda Cancela García	3372654

Y considerando que reúne los requisitos establecidos en la Opción: **TITULACIÓN INTEGRAL:**
TESIS PROFESIONAL "EXTRACCIÓN DE SILUETAS DEL CAMINAR DE UNA PERSONA", damos
a usted nuestra autorización para que proceda a imprimir su trabajo profesional

Sin otro particular, quedo de usted.

ATENTAMENTE

Excelencia en Educación Tecnológica®


M.C. CARLOS RAFAEL LIZARRAGA ARREOLA
JEFE DE LA DIVISIÓN DE ESTUDIOS PROFESIONALES



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

INSTITUTO TECNOLÓGICO DE CULIACÁN

DEPARTAMENTO DE DIVISIÓN DE
ESTUDIOS PROFESIONALES



Juan de Dios Bätz 310 Pte. Col. Guadalupe C.P.80220 Culiacán, Sinaloa. Tel. 667-454-0100
y 667-713-3804 | <https://www.culiacan.tecnm.mx>



2022 Flores
Año de
Magón
PRESENCIA DE LA REVOLUCIÓN MEXICANA

Dedicatoria

Hago uso de este espacio para dedicar la culminación de 5 años de estudios a mi madre, por sus constantes sacrificios en busca de que sus hijos vivieran una vida plena, una vida que tal vez no tuvo, pues sé que mi educación y todos los frutos que vengan de ella estarán sostenidos por los cimientos de su amor como madre. A mis hermanos, Roberto y Quetzally, quienes compartieron las adversidades de la vida, que compartieron el pan cuando no había, que siempre tenían dinero, tiempo y espacio para su hermano menor, sé que, sin sus consejos y amor jamás hubiera podido alcanzar mis metas, los cimientos de este documento están en el apoyo de ustedes tres, Quetzally dándome su hogar, Roberto su guía y mi madre Alma su corazón.

A mis amigos cercanos que siempre tuvieron un hombro para que me recargara y una mano para ayudarme, que va desde José Manuel que me prestaba dinero para mi pasaje en la preparatoria, Alfonso que me mostro que ser de los mejores es cuestión de acciones e influencio mi proceso como ingeniero, especialmente a Graciela que nunca dejo de prestarme lápices desde la preparatoria, y a todos aquellos que estoy olvidando, gracias.

Aun cuando mis pasos llevan la huella del esfuerzo de mi familia y amigos, también estoy arraigado fuertemente a distintas personas de quienes fui aprendiendo en distintas instituciones educativas, agradezco al maestro Damaso, por sembrar la semilla para que fuera una persona de principios, a la maestra norma por hacerme explotar mis capacidades como estudiante, al maestro Gerardo por hacerme creer en mí, no sentir vergüenza por mi amor a las matemáticas y consecuentemente explorar una carrera de ingeniería, al maestro Rojo Lugo, por no dejar que me rindiera y hacerme ver que mis problemas eran minúsculos, y al resto de maestros que vieron en mi un estudiante con potencial pero con problemas y decidieron ayudarme, gracias.

Dejo estos últimos 2 párrafos para dos maestros que marcaron especialmente mi camino en ingeniería.

Agradezco de todo corazón al doctor Héctor Rodríguez Rangel, de quien estoy consciente que tal vez jamás pude llenar la expectativa que tenía de mi como estudiante, pero que jamás dudo en abrir espacio en su vida para mis problemas personales, quien pudo haber actuado como una persona no racional y no apoyarme en las situaciones en las que me encontraba enredado, probablemente jamás hubiera conocido el placer de graduarme si no fuera por su guía, consejos y ayuda, que probablemente fuera un hombre distinto de no haberlo conocido, la grandeza nace del caos, lo llevare en mi alma en cada paso que tome en mi carrera, como el hombre que hizo que fuera un científico de datos, gracias doctor.

Y por último a la maestra que me enseñó que el trabajo duro es la única manera de hacer las cosas, a no poner excusas y aceptar mis errores, a buscar a ser mejor persona, mejor estudiante, verdaderamente la persona que me enseñó lo que es ser un ingeniero de verdad, por quien probablemente hice todas las líneas de Código que escribí y estoy por escribir, tal vez de no haberla conocido pudiera haberme graduado, titulado y todas las cosas en el mundo, pero tengo la certeza de que no hubiera sido bueno en nada que me hubiese propuesto. Jamás pude cumplir mi promesa de hacer las cosas a tiempo, en su forma y sin problemas, pero, aun así, aprendí tanto de la maestra Ekatherine que me cuesta trabajo saber que es mío y que es de ella cuando alguien me pregunta sobre un objeto. Hablo con total seguridad de que si no deje mi carrera y empecé a buscar a resolver los problemas que la vida me ha puesto fue una influencia directa a sus clases de programación orientada a objetos, jamás hubiera hecho mi servicio social si no me hubieran dado la oportunidad de dar asesorías, ni mis residencias y mucho menos este documento, por eso y mil cosas que estoy olvidando pero que llevo en mi ser, muchísimas gracias, maestra Katy, solo algún poder superior es capaz de saber todo lo que hizo por mí. Espero que mi trabajo de vida sea digno del que usted y el doctor han hecho.

Extracción de siluetas del caminar de una persona

Resumen

El reconocimiento del andar humano es una característica biométrica que ha despertado el interés de la comunidad científica en los últimos años. La biometría son las medidas biológicas, o características físicas, que se pueden utilizar para identificar a las personas.

Uno de los métodos principales para el reconocimiento del andar humano es basado en siluetas, que a su vez generan imágenes de energía que representan a la persona a través de su ciclo de caminata. El perfil o contorno que presenta una figura se le denomina con el término de silueta.

En el presente trabajo se desarrolla un sistema, que haciendo usos de técnicas de computación visual, es capaz de extraer las siluetas del ciclo de caminata de los usuarios y generar sus respectivas imágenes de energía, que tienen como propósito, representar el ciclo de caminata particular de una persona a través de una imagen.

El sistema se compone de dos etapas, la primera es la extracción de las siluetas de personas caminando en un video y la segunda, el cálculo del promedio de las siluetas obtenidas para generar una imagen de energía del ciclo del caminado.

Para probar el sistema, se utiliza la base de datos CASIA-B que es una de las principales disponibles en el área del reconocimiento del andar humano debido a su gran tamaño y calidad de los videos.

La base de datos generada en esta propuesta funcionaria como datos de entrada para probar modelos predictivos capaces de identificar a una persona a partir de su imagen de energía.

Palabras clave: Característica biométrica, Siluetas, Imágenes de energía, CASIA-B, Modelos predictivos, Identificación biométrica, Andar humano

Contenido

1. INTRODUCCIÓN.....	5
1.1 Antecedentes.....	7
1.2 Descripción del problema	10
1.3 Hipótesis o preguntas de investigación.....	12
1.4 Objetivo general.....	12
1.5 Objetivos específicos	12
1.6 Justificación	13
1.7 Alcance.....	14
2. Fundamento teórico.....	15
2.1 Herramientas y Tecnologías	15
2.1.1 Computación Visual	15
2.1.2 Umbralizacion	15
2.1.3 CUDA.....	16
2.1.4 Python.....	17
2.1.5 OpenCV	17
2.1.6 Tensorflow	17
2.1.7 Operaciones morfológicas.....	18
2.1.8 Eliminacion Guassiana.....	18
2.2 Conceptos de Inteligencia artificial:.....	18
2.2.1 Inteligencia artificial.....	18
2.2.2 Técnicas de inteligencia artificial	18
2.2.3 Parámetros biométricos	19
2.2.4 Red Neuronal.....	19
2.2.5 Redes neuronales Convolucionales	19

2.2.6 Aprendizaje no supervisado	19
2.2.7 Aprendizaje supervisado	20
2.2.8 Aprendizaje maquina	20
2.2.9 Aprendizaje Profundo o Deep learning.....	21
2.2.10 Elementos clave del machine learning:	22
2.2.11 Aprendizaje inductivo:	22
3. Procedimiento	24
3.1 Investigación y practica.....	24
3.1.1 Trabajos previos.....	24
3.1.2 Talleres de practica.....	24
3.1.3 Computación visual y manipulación de imágenes:	24
3.1.4 Siluetas, imágenes de energía y operaciones morfológicas	25
3.2 Experimentación y datos.....	26
3.2.1 Datos a utilizar	26
3.2.2 Prueba y validación de algoritmos.....	27
3.2.3 Diseño de entrada y salida.....	29
3.3 Manipulación de datos y extracción de siluetas	30
3.3.1 Descomposición de videos:.....	30
3.3.2 Transformación de videos por medio de operaciones morfológicas y gaussianas:.....	30
3.4 Imágenes de energía.....	34
3.4.1 Extracción de silueta de la imagen.....	34
4. Resultados	39
5. Conclusiones.....	42
6. Competencias desarrolladas	43

7. Fuentes de información.....	44
--------------------------------	----

LISTA DE FIGURAS

Figura 1 Proceso de umbralizacion o thresholding	16
Figura 2 Logotipo del lenguaje Python	17
Figura 3 Creación de una imagen de energía.....	24
Figura 4 Video del andar humano de la base de datos CasiaB	26
Figura 5 Uso del método de Otsu	¡Error! Marcador no definido.
Figura 6 Imagen con silueta resaltada utilizando un tipo de Gauss	¡Error! Marcador no definido.
Figura 7 Otro tipo de imagen resultante del método de Gauss para la de extracción de fondos	¡Error! Marcador no definido.
Figura 8 Funcionamiento general	30
Figura 9 Silueta resultante.....	31
Figura 10 Silueta obtenida.....	34
Figura 11 Promediado de siluetas imagen de energia.....	35
Figura 12 Preprocesamiento de información y creación de imagen de energía	41
Figura 13 Tipo de imágenes de energía obtenidas.....	41

Lista de Algoritmos

Algoritmo 1	Extraccion de fondos y resaltado de siluetas	33
Algoritmo 2	Obtención de masa central de una imagen	36
Algoritmo 3	Obtención de siluetas.....	37
Algoritmo 4	Creación de imagen de energía.... ¡Error! Marcador no definido.	

Lista de tablas

Tabla 1	Resultados obtenidos de generación de GEI por cada grupo de la base de datos CASIA-B.....	40
----------------	---	----

“Extracción de siluetas del caminar de una persona”

1. INTRODUCCIÓN

Desde la década del 2010, los niveles de incidencia delictiva se han mantenido altos [24]. Los asaltos, robos a casa-habitación, secuestros y asesinatos son de los delitos que alarman a la población general debido a la gravedad y cotidianidad. Las instituciones, tanto públicas como privadas, necesitan salvaguardar sus bienes de personas no autorizadas, que puedan realizar algún tipo de robo dentro de sus instalaciones.

Las cámaras de vigilancia presentan cierto nivel de seguridad, pero tienen poca efectividad por si solas, ya que es necesario personal adicional encargado del monitoreo.

El reconocimiento del andar humano es una técnica biométrica no invasiva, que se encarga de identificar a una persona por medio de su forma de caminar. En esta técnica se procesan imágenes de un video para obtener datos que reflejen la identidad del sujeto que está caminando.

En los últimos años, varios investigadores han dirigido sus trabajos sobre esta característica biométrica, debido a las ventajas que ofrece, como ejecutar el reconocimiento a distancia e incluso sin el consentimiento del sujeto que está caminando, evitando así el uso de una técnica invasiva que tendría consecuencias legales.

El reconocimiento del andar humano mediante imágenes de energía de la caminata representa una alternativa no invasiva y, además, es una de las técnicas mejor exploradas en la identificación biométrica. Estas imágenes representan el ciclo de caminata de una persona y se obtienen calculando el promedio de las siluetas extraídas de un video.

La contribución de este trabajo es la generación de un sistema que utiliza técnicas de computación visual para manipular videos, en donde exista un sujeto caminando, para la extracción de su silueta y su consecuente creación de imagen de energía. La técnica de extracción de las siluetas está basada en aprendizaje profundo, utilizando una red neuronal convolucional llamada ResNet101 que se encarga de detectar y extraer las siluetas humanas de los videos cuadro por cuadro.

Las Redes neuronales convolucionales son un tipo de redes neuronales artificiales donde existe una abstracción a una neurona que corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria de un cerebro biológico.

Para la implementación de este tipo de técnicas es necesario un gran procesamiento de extracción de características, las cuales hacen a una persona diferente de otra. Donde a partir de estas características es posible identificar correctamente a una persona.

Existen diversas formas de extraer las características de una persona para diferenciarla por su manera de caminar (andar) puede ser:

- Por su silueta.
- Por imágenes de energía.
- Generando un esqueleto,

Por una representación 3D.

Sin embargo, estos métodos pueden ser complicados de implementar.

Entre los factores que pueden afectar en la identificación del andar se encuentran:

- Los ángulos de vista pueden variar dependiendo de la posición de la cámara,
- La visibilidad afectada por la distancia entre el usuario y la cámara,
- Si el usuario se encuentra en la lejanía, perjudica la detección,
- Si el usuario se encuentra demasiado cerca puede que no se logre captar correctamente la silueta; solo por mencionar algunos casos.

En este proyecto se explorará determinar algún mecanismo para extraer las características únicas de una persona al caminar, con la finalidad de ser usadas para un sistema de videovigilancia que restringe el acceso a ciertas zonas.

Para definir la metodología de desarrollo del proyecto, se divide en 3 etapas. Estas etapas son las siguientes:

- Modelado, diseño y codificación del sistema.
- Evaluación y mantenimientos al sistema y análisis de resultados.
- Documentación del proyecto

Palabras clave: Silueta, imágenes de energía, biometría, identificación biométrica, técnica no invasiva, inseguridad, red neuronal

1.1 Antecedentes

El identificar entidades según su apariencia por medio de la computación, es un tópico con diversos enfoques, el enfoque en la apariencia o sin patrones han sido propuestos desde la década de los 90s [5], la origen de estos enfoques es debido a que los métodos usuales para el reconocimiento de objetos no suelen producir una descripción viable ni compacta de las imágenes, en donde el contenido debe ser privado de información sobre el fondo y ocultar otro tipo de información no vital para la detección [5].

Cunado D. introdujo por primera vez el modelado de movimiento utilizando la Transformada de Velocidad Huff para extraer la velocidad angular de la articulación de la cadera modelando la marcha humana como un péndulo en movimiento. Se logró una tasa de reconocimiento del 90 % utilizando la firma derivada en un conjunto de datos que contiene 10 sujetos [12].

”Automated person recognition by walking and running via model-based approaches. Pattern Recognition” simuló la marcha humana como un oscilador acoplado dinámico utilizado para extraer movimientos angulares de caderas y rodillas mediante la recopilación de pruebas. Este método se probó en un pequeño conjunto de datos de 20 peatones y corredores con una tasa de reconocimiento del 91 % basada en la firma de la marcha obtenida del análisis de Fourier del movimiento angular [13].

La imagen de energía de la marcha (GEI) es una representación básica basada en la silueta presentada por ”Individual recognition using gait energy image” en la que la firma de la caminata se construye tomando el promedio de siluetas para un ciclo de pasos. Los resultados experimentales confirmaron que se pueden alcanzar tasas de reconocimiento más altas para alcanzar el 94,24 % para un conjunto de datos de 3141 sujetos. Sin embargo, dicho método no funciona bien al cambiar la apariencia [7].

En ”Gait Recognition Using a View Transformation Model in the Frequency Domain” se propone el uso de una técnica de reconocimiento del andar humano utilizando el volumen de siluetas extraídas durante la caminata. Con una cámara infrarroja, se extraen las siluetas de los sujetos caminando en distintos ángulos de visión. (GSV, Gait Silhouette Volume.) La GSV es generada haciendo uso de las rebanadas superiores, céntricas e inferiores de las siluetas obtenidas. El vector de características GSV resultante, es clasificado mediante un modelo de transformación de ángulo de visión. Este modelo tiene la capacidad de clasificar las GSV en diferentes combinaciones de ángulos debido a que su proceso algorítmico interno normaliza todos los ángulos a uno generalizado para todas las vistas posibles [23].

Zhao, G. usó un algoritmo de optimización local para construir un modelo 3D a partir de imágenes de video filmadas por múltiples cámaras. Las funciones de clasificaciones derivadas de los segmentos corporales incluyen parámetros estáticos y funciones dinámicas de las trayectorias de movimiento de las extremidades [17].

La imagen de silueta en movimiento (MSI) es una representación similar a la GEI propuesta por Lam and Lee donde la intensidad de cada píxel se calcula en función del historial temporal de movimiento de los píxeles correspondientes a lo largo de un ciclo de caminata completo. Los experimentos realizados con el gran conjunto de datos de caminata de SOTON mostraron que se puede lograr el 87 % de Efectividad [8].

Bashir K. utilizó el campo de flujo óptico denso calculado mediante el método de Brox para cada cuadro de todo el ciclo de la forma de caminar para extraer cuatro tipos diferentes de descriptores de movimiento (imagen de intensidad de movimiento y cuatro imágenes de dirección de movimiento) basados en los componentes y la magnitud del flujo óptico horizontal y vertical, sus experimentos propuestos en los conjuntos de datos de caminata de CASIA y SOTON con las variables de ropa y bolsas superan los estudios descritos anteriormente [9].

Recientemente, Yam. C., propuso un enfoque que combina similitudes parciales con el enfoque, basado en el supuesto de que los objetos 3D comparten el mismo plano de visualización en diferentes vistas. El modelo paramétrico 3D del cuerpo humano es una deformación de la postura y la forma es transformada por el modelo del modelo utilizando los datos de silueta 2D.

Los modelos de movimiento representan la cinemática o dinámica del cuerpo humano o sus diferentes partes a lo largo del tiempo. El modelo de movimiento utiliza varias restricciones que apoyan el proceso de extracción, como el radio máximo de giro de las extremidades inferiores [6].

Al necesitar una resolución para el reconocimiento en un ambiente en el que generalmente hay cambios constantes en las imágenes es necesario utilizar otro tipo de enfoque.

Los enfoques basados en apariencia para el reconocimiento de la forma de caminar no requieren la computación previa de los patrones de la marcha o el andar, los rasgos son extraídos de todo el cuerpo sin necesidad de extraer explícitamente cada parte del cuerpo humano. La mayoría de los enfoques de este tipo son basados en datos obtenidos directamente de la silueta excluyendo el fondo de la imagen. Además, es común que se apliquen técnicas avanzadas de aprendizaje automático, como redes profundas neuronales o maquinas vectoriales de soporte de capas múltiples [6].

"Incremental learning for video-based gait recognition with LBP flow" describió un marco de aprendizaje incremental que utiliza el flujo óptico para el reconocimiento de la marcha. El operador de patrón binario local (LBP) se utiliza para codificar la información de textura del flujo óptico [11].

Para el enfoque basado en modelos, se establece un modelo previo para hacer coincidir las imágenes reales con este modelo predefinido y, por lo tanto, extraer las características de la marcha correspondientes una vez que se obtiene la mejor coincidencia. Por lo general, cada cuadro que contiene un sujeto que camina se ajusta a un modelo temporal o espacial previo para extraer explícitamente características de la marcha como la distancia de la zancada, las medidas angulares, las trayectorias de las articulaciones o las medidas antropométricas [6].

"A new approach for multi-view gait recognition on unconstrained paths. Journal Of Visual Communication And Image Representation" utilizó datos 3D inclinados para crear una firma biométrica de marcha inmutable para la visibilidad de los peatones en carreteras ilimitadas. Una máquina de vectores de soporte se utilizó en el proceso de clasificación de dos conjuntos de datos que contienen un total de 62 objetos.

Kastaniotis, D., utilizó el sensor Kinect de Microsoft para identificar a una persona a través de un proceso de estimación de la postura que recupera el esqueleto de un peatón. Sobre la base de un conjunto de datos públicos de 30 personas, se informa una alta tasa de sensibilización del 93,3

En el trabajo de "A model-based gait recognition method with body pose and human prior knowledge. Pattern Recognition" se desarrolla un sistema de reconocimiento del andar humano basado en modelos. En él se destacan la pose, el movimiento, los ángulos de inclinación y longitud de las partes del cuerpo como las características a obtener de cada sujeto de prueba. Después, los vectores de características se fusionan para generar un solo vector de características. Posteriormente se utiliza una red neuronal convolucional (CNN) para clasificar este vector generado y obtener los resultados del reconocimiento del andar humano.

1.2 Descripción del problema

Tradicionalmente, la detección a violaciones a zonas restringidas se soluciona con el uso de personal adicional para que se encuentre físicamente, vigilando alguna inconsistencia al acceso de la institucional. Sin embargo, esta acción no es del todo confiable pues se le da una total confianza a una persona. Además, el coste que implica la contratación de esta persona. En este trabajo se pretende realizar un sistema de extracción de características del caminar de una persona con la finalidad de ser aplicada en un sistema de video vigilancia de zonas restringidas.

El uso de contraseñas, tarjetas inteligentes, llaves o una multitud de elementos intercambiables no es práctico y no proporciona la seguridad necesaria para permitir que personas no deseadas accedan a una ubicación en particular. Se debe instalar una cerradura biométrica en la puerta de esta área para monitorizar la actividad humana que desea ingresar al área restringida y asegurarse de que sea la persona deseada.

Debido a lo anterior, junto a la rápida evolución de técnicas en la manipulación de imágenes y técnicas de computación visual, han surgido diferentes técnicas de reconocimiento a través de características biométricas que tienen altos estándares de calidad y son utilizadas ampliamente en el mercado.

Pero la mayoría de estas técnicas resultan invasivas, requieren de la colaboración del usuario, y necesitan una cooperación directa por parte de la persona para ser identificada. Colocar el dedo en un escáner de huellas digitales o mostrar el rostro frontalmente en una cámara de video son dos ejemplos de lecturas biométricas que requieren de la autorización y participación por parte del usuario.

Es por esto por lo que es preciso diseñar sistemas de reconocimiento biométricos no invasivos y de alta calidad que brinden altos porcentajes de éxito en sus lecturas. Su implementación en hogares, aeropuertos, hospitales, bancos o cualquier área restringida para personal no autorizado, aportaría un sistema de alertas de seguridad contra amenazas y que no necesitan ser operados o atendidos por personal humano.

El identificar a una persona por medio de su andar representa resolver grandes retos. Entre uno de los grandes retos se encuentra el determinar las características que hacen única a una persona por medio de su caminar. A partir de estas características es posible identificar a una persona mediante el uso de modelos de aprendizaje de inteligencia artificial como son las redes neuronales artificiales.

Al solucionar el problema mediante el uso de cámaras de video, se reducen los gastos de incrementar al personal de seguridad y se garantiza la seguridad de una zona restringida, que regularmente es una zona crítica para la institución.

La identificación de una persona por medio de su andar es un trabajo explorado. Donde una parte crítica para resolver el problema, es la extracción de la información biométrica de la persona, que es necesaria para realizar la identificación.

1.3 Hipótesis o preguntas de investigación

Las técnicas de visión computacional permiten extraer características importantes y únicas que permiten la identificación de una persona mediante su andar a partir de secuencia de imágenes.

1.4 Objetivo general

Generar una herramienta capaz de extraer las características del andar de una persona a partir de secuencia de imágenes utilizando técnicas de computación visual.

1.5 Objetivos específicos

- Investigar los métodos de visión computacional para la extracción de características que permiten identificar una persona por medio de su andar.
- Desarrollar un sistema capaz de procesar la información contenida en videos y descomponerla en secuencias de imágenes por medio de técnicas de visión computacional.
- Desarrollar un sistema capaz de extraer las características del andar de una secuencia de imágenes.
- Implementar ajustes dentro del sistema para el procesamiento y la extracción cumpla una tasa de éxito superior al 90%.

1.6 Justificación

En la actualidad, las instituciones privadas o públicas están obligadas a mantener intacta la integridad de sus empleados o clientes dentro de la misma. A su vez, existen personas que atentan ante la integridad de las instituciones para obtener un beneficio ilícito o afectar negativamente a la institución. Muchas instituciones confían en métodos convencionales, como las cerraduras, pero estas ofrecen poco nivel de seguridad y practicidad. Uno de los problemas principales de los sistemas de acceso tradicionales se encuentra en que las llaves o tarjetas de acceso son objetos que físicamente pueden perderse o duplicarse.

En caso de contraseñas o pines, estos pueden ser olvidados o ser robados por personas no autorizadas. Esto lleva a que esa misma institución aplique medidas de seguridad cada vez mayores y mejores para asegurar la integridad de sus empleados, clientes, o bienes. Una de estas medidas consiste en la aplicación de sistemas de seguridad biométricas para la vigilancia y el acceso a zonas de acceso restringido.

Una solución para resguardar las zonas restringidas de las instituciones o empresas es la contratación y personal extra con la finalidad de monitorear las zonas restringidas o cruciales. Sin embargo, al ser una persona puede no ser completamente confiable.

Debido a que puede no estar atento siempre y puede no llegar a detectar correctamente un intruso, además que, representa un gasto constante extra, puede ser corrompible, etc.

Por lo que un sistema de monitoreo de personas no autorizadas dará una mayor seguridad a la zona restringida a un costo menor de tener personal extra-encargado.

1.7 Alcance

Se utilizarán diversas técnicas de visión computacional para extraer la información contenida en videos y descomponerla en una serie de fotogramas, además se utilizará como punto de partida para el diseño del proyecto la base de datos CASIAB.

Con la información recabada se construirán imágenes de energía que puedan ser utilizadas para la clasificación e identificación del andar humano.

Como recursos disponibles para la computación de este proyecto se tiene disponible un equipo con un procesador Intel Core i7 5.1 GHz con GPU Nvidia GeForce GTX 1050 TI y 16 gb de RAM Hyperx DDR4.

El entregable final es la herramienta extractora de características para el identificar del andar humano.

Este proyecto no abarca el proceso de entrenamiento y refinamiento de un modelo biométrico identificador de personas caminando.

2. Fundamento teórico

2.1 Herramientas y Tecnologías:

2.1.1 Computación Visual

La computación visual es un campo de la computación que se ocupa de la adquisición, análisis y síntesis de datos visuales mediante el uso de recursos informáticos. Abarca varios campos de la ciencia (informática, en particular), matemáticas, física y ciencias cognitivas. La informática visual tiene como objetivo permitirnos controlar e interactuar con actividades a través de la manipulación de imágenes visuales, ya sea como objetos directos o, simplemente, como representaciones de objetos no visuales [1].

2.1.2 Umbralizacion

La umbralización es uno de los más importantes métodos de segmentación. El objetivo es convertir una imagen en escala de grises a una nueva con sólo dos niveles, de manera que los objetos queden separados del fondo. El histograma de la Figura corresponde a una imagen compuesta de objetos claros sobre un fondo oscuro. Una forma de extraer el objeto del fondo es seleccionar un umbral de intensidad T que separe los dos conjuntos. De este modo, cualquier punto (x, y) para el que $f(x, y) > T$, *pertenecerá* al objeto, en caso contrario, *pertenecerá* al fondo. Si los valores de gris del objeto y del resto de la imagen difieren claramente, entonces el histograma mostrará una distribución bimodal, con dos máximos distintos, separados por una región vacía. Con lo cual se logrará una separación perfecta entre el objeto y el fondo, al establecer un valor umbral ubicado en esta región del histograma [2].

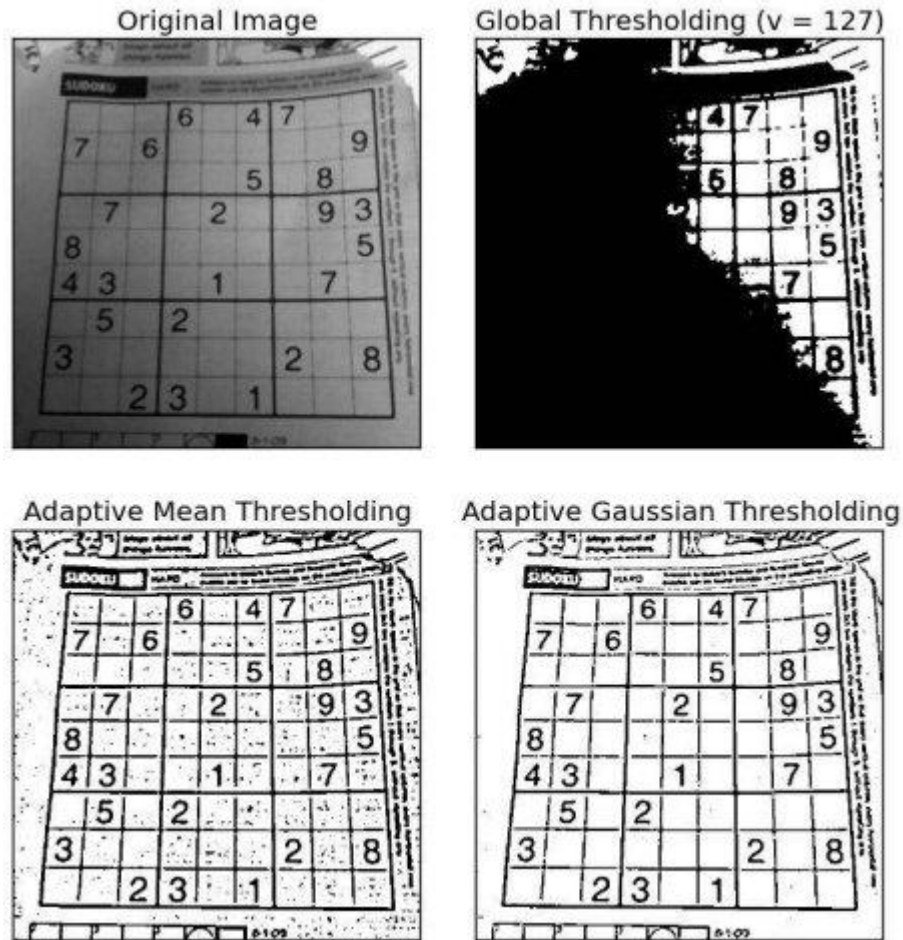


Figura 1 Proceso de umbralizacion o thresholding

2.1.3 CUDA

CUDA es un acrónimo de Compute Unified Device Architecture y hace referencia a la plataforma de GPGPU de NVIDIA. Esto significa que con CUDA nos referimos a la ejecución no solo de programas relacionados con la generación de gráficos en los juegos, sino de propósito general, lo que permite utilizar la alta potencia de cálculo de las tarjetas gráficas para acelerar ciertas tareas en paralelo donde una CPU convencional es menos eficiente.

Existe desde hace años un ecosistema de aplicaciones y herramientas alrededor de CUDA centrado en el mundo científico y la ingeniería y en ramas distintas de cada uno. Desde la medicina hasta en el diseño de automóviles. Lo cual le permitió crecer a NVIDIA más allá del hardware para juegos de PC y ampliar su cuota de mercado potencial.

CUDA es más bien es una filosofía a la hora de programar algoritmos que se ejecutarán en una GPU de NVIDIA, aunque también hay posibilidades de hacerlo en un procesador central e incluso en un chip de la competencia. En la actualidad existen diversos lenguajes de programación que tienen las extensiones CUDA correspondientes. Entre los que se incluyen: C, C++, Fortran, Python y MATLAB [3].

2.1.4 Python

Python es un lenguaje de programación orientado a objetos de alto nivel e fácil de interpretar con una sintaxis fácil de leer. Python, ideal para prototipos y tareas ad hoc, tiene un amplio uso en la informática científica, el desarrollo web y la automatización. Como lenguaje de programación para principiantes y con fines generales, Python admite muchos de los mejores científicos de computadoras y desarrolladores de aplicaciones a nivel global [4].

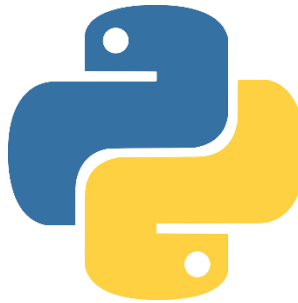


Figura 2 Logotipo del lenguaje Python

2.1.5 OpenCV

Una biblioteca de código abierto que contiene implementaciones que abarcan más de 2500 algoritmos. Además, está especializada en el sistema de visión artificial y machine learning [10].

2.1.6 Tensorflow

TensorFlow es una biblioteca de código abierto dirigida al aprendizaje automático a través de una serie de tareas. Ha sido desarrollado por Google para satisfacer las necesidades de sistemas capaces de construir y entrenar redes neuronales para detectar y descifrar patrones y correlaciones, análogos al aprendizaje y razonamiento usados por los humanos. Actualmente es utilizado tanto para la investigación como para la producción de productos de Google, remplazando el rol de su predecesor de código cerrado, DistBelief. TensorFlow fue originalmente desarrollado por el equipo de Google Brain para uso interno en Google antes de ser publicado bajo la licencia de código abierto Apache 2.0 el 9 de noviembre de 2015 [26].

2.1.7 Operaciones morfológicas

Las operaciones morfológicas son un conjunto amplio de operaciones que sirven para procesar imágenes basadas en formas geométricas. En una operación morfológica, cada píxel de la imagen se ajusta en función del valor de otros píxeles de su entorno. Si se seleccionan la forma y el tamaño del entorno adecuadamente, se puede definir una operación morfológica sensible a ciertas formas de la imagen de entrada [27].

2.1.8 Eliminación Gaussiana

El algoritmo de eliminación Gaussiana realiza lo que se llama una reducción de filas en una matriz donde se realizan una serie de operaciones sobre los valores hasta que la esquina izquierda inferior de la matriz consiste solamente en ceros [28].

Existen tres tipos de operaciones elementales de filas

- Intercambiar dos filas
- Multiplicar una fila por un escalar diferente de cero
- Sumar el múltiplo de una fila a otra fila

2.2 Conceptos de Inteligencia artificial:

2.2.1 Inteligencia artificial

La inteligencia artificial general, a veces denominada “inteligencia artificial fuerte” o “inteligencia artificial de nivel humano”, hace referencia a la capacidad de un sistema informático de superar a las personas en cualquier tarea intelectual. Es el tipo de inteligencia artificial que se ve en las películas en las que los robots tienen pensamientos conscientes y actúan según sus propios motivos.

En teoría, un sistema informático que ha conseguido inteligencia artificial general podría resolver problemas sumamente complejos, emitir juicios en situaciones inciertas e incorporar conocimientos previos a su razonamiento actual. Podría tener creatividad e imaginación a la par que las personas y podría realizar muchísimas más tareas que la inteligencia artificial estrecha [32].

2.2.2 Técnicas de inteligencia artificial

Las técnicas de inteligencia artificial hacen referencia a diferentes campos de investigación y de desarrollo de aplicación de la inteligencia artificial. La IA es el tipo de inteligencia que muestra los distintos artefactos creados por los humanos, normalmente referido a los sistemas informáticos. Pero también hace referencia al ámbito de la investigación científica en la intención de crear entornos que hagan uso de dicha inteligencia [18].

2.2.3 Parámetros biométricos

- FAR (*False Acceptance Rate*): tasa de falso positivo, hace referencia a la probabilidad de que un usuario no autorizado sea aceptado. Este parámetro deberá ajustarse para evitar el fraude en los sistemas biométricos.
- FRR (*False Rejection Rate*): tasa de rechazo erróneo, la probabilidad de que un usuario que está autorizado sea rechazado a la hora de intentar acceder al sistema. Si los usuarios son rechazados erróneamente con frecuencia, parecerá que el sistema no funciona correctamente y deberá ser revisado.
- FER (*Failure to Enroll Rate*): tasa de fallo de alistamiento, hace referencia a los usuarios que son rechazados cuando van a ser registrados a causa de la mala calidad de su muestra.
- FTE (*Failure to Enroll*): indica la probabilidad numérica de que alguien no sea registrado a causa de un fallo a la hora de crear un patrón [29].

2.2.4 Red Neuronal

Una red neuronal es un método de la inteligencia artificial que enseña a las computadoras a procesar datos de una manera que está inspirada en la forma en que lo hace el cerebro humano. Se trata de un tipo de proceso de *machine learning* llamado aprendizaje profundo, que utiliza los nodos o las neuronas interconectados en una estructura de capas que se parece al cerebro humano. Crea un sistema adaptable que las computadoras utilizan para aprender de sus errores y mejorar continuamente. De esta forma, las redes neuronales artificiales intentan resolver problemas complicados, como la realización de resúmenes de documentos o el reconocimiento de rostros, con mayor precisión [19].

2.2.5 Redes neuronales Convolucionales

Las Redes neuronales convolucionales son un tipo de redes neuronales artificiales en donde las «neuronas» corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria de un cerebro biológico. Este tipo de red es una variación de un perceptrón multicapa, sin embargo, debido a que su aplicación es realizada en matrices bidimensionales, son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones [20].

2.2.6 Aprendizaje no supervisado

El aprendizaje no supervisado es una de las formas en que Machine Learning (ML) "aprende" los datos. El aprendizaje no supervisado tiene datos sin etiquetar que el algoritmo tiene que intentar entender por sí mismo. El aprendizaje supervisado es en el que se etiquetan los conjuntos de datos para que haya una clave de respuestas con

la que la máquina pueda medir su precisión. El objetivo es simplemente dejar que la máquina aprenda sin ayuda o indicaciones de los científicos de datos. En el camino, también deberá aprender a ajustar los resultados y agrupaciones cuando haya resultados más adecuados, permitiendo que la máquina comprenda los datos y los procese como mejor le parezca. El aprendizaje no supervisado se utiliza para explorar datos desconocidos. Puede revelar patrones que podrían haberse pasado por alto o examinar grandes conjuntos de datos que serían demasiado para que los abordara una sola persona [30].

2.2.7 Aprendizaje supervisado

El aprendizaje supervisado es el tipo de algoritmo de Machine Learning más frecuente. Utiliza un conjunto de datos conocidos (denominado conjunto de datos de entrenamiento) para entrenar un algoritmo con un conjunto de datos de entrada conocidos (denominados características) y respuestas conocidas para realizar predicciones. El conjunto de datos de entrenamiento incluye datos de entrada etiquetados que se emparejan con los valores de salida o de respuesta deseados. A partir de él, el algoritmo de aprendizaje supervisado intenta crear un modelo estableciendo relaciones entre las características y los datos de salida para realizar predicciones acerca de los valores de respuesta para un nuevo conjunto de datos.

Antes de aplicar el aprendizaje supervisado, el aprendizaje no supervisado se suele utilizar primero para encontrar patrones en los datos de entrada que indiquen características potenciales; luego, la ingeniería de características las transforma para que resulten más adecuadas para el aprendizaje supervisado.

Además de identificar las características, también se debe identificar la categoría o la respuesta correctas para todas las observaciones del conjunto de datos de entrenamiento, que es un paso muy laborioso. Con el aprendizaje semi supervisado, es posible entrenar modelos con tan solo algunos datos etiquetados, lo que ayuda a reducir el trabajo de etiquetado. Una vez entrenado el algoritmo, se suele utilizar un conjunto de datos de prueba, que no se haya utilizado en el entrenamiento, para predecir el rendimiento y validar el algoritmo. Para obtener resultados de rendimiento precisos, es fundamental que el conjunto de datos de entrenamiento y el conjunto de datos de prueba sean correctos; es decir, que tanto los datos del modelo como los del entorno de producción se hayan validado correctamente [31].

2.2.8 Aprendizaje maquina

El aprendizaje maquina es una forma de la IA que permite a un sistema aprender de los datos en lugar de aprender mediante la programación explícita. Sin embargo, machine learning no es un proceso sencillo. Conforme el algoritmo ingiere datos de entrenamiento, es posible producir modelos más precisos basados en datos. Un

modelo de machine learning es la salida de información que se genera cuando entrena su algoritmo de machine learning con datos. Después del entrenamiento, al proporcionar un modelo con una entrada, se le dará una salida. Por ejemplo, un algoritmo predictivo creará un modelo predictivo. A continuación, cuando proporcione el modelo predictivo con datos, recibirá un pronóstico basado en los datos que entrenaron al modelo.

El Machine learning permite modelos a entrenar con conjuntos de datos antes de ser implementados. Algunos modelos de machine learning están online y son continuos. Este proceso iterativo de modelos online conduce a una mejora en los tipos de asociaciones hechas entre los elementos de datos. Debido a su complejidad y tamaño, estos patrones y asociaciones podrían haber sido fácilmente pasados por alto por la observación humana. Después de que un modelo ha sido entrenado, se puede utilizar en tiempo real para aprender de los datos. Las mejoras en la precisión son el resultado del proceso de entrenamiento y la automatización que forman parte del machine learning.

Las técnicas de machine learning son necesarias para mejorar la precisión de los modelos predictivos. Dependiendo de la naturaleza del problema empresarial que se está atendiendo, existen diferentes enfoques basados en el tipo y volumen de los datos. En esta sección, discutimos las categorías del machine learning [25].

2.2.9 Aprendizaje Profundo o Deep learning

Deep learning es un subconjunto de machine learning (que a su vez es parte de la inteligencia artificial) donde las redes neuronales, algoritmos inspirados en cómo funciona el cerebro humano, aprenden de grandes cantidades de datos.

Los algoritmos de deep learning realizan una tarea repetitiva que ayuda a mejorar de manera gradual el resultado a través de “deep layers” lo que permite el aprendizaje progresivo. Este proceso forma parte de una familia más amplia de métodos de machine learning basados en redes neuronales.

Deep learning ha tenido un gran impacto en todas las industrias. Por ejemplo, en las ciencias de la vida, el aprendizaje profundo se puede utilizar para el análisis avanzado de imágenes, la investigación, el descubrimiento de medicinas, la predicción de problemas de salud, así como síntomas de enfermedades, y la aceleración de conocimientos a partir de la secuenciación genómica. En el transporte, puede ayudar a los vehículos autónomos a adaptarse a las condiciones cambiantes; y a su vez ser utilizado para proteger infraestructuras críticas.

Existe una idea errónea, la cual ha tomado fuerza con el tiempo, de que deep learning es una tecnología que compite en contra de machine learning. Para poner las cosas

en perspectiva, deep learning es un subdominio de machine learning; con un poder computacional acelerado y grandes conjuntos de datos. Los algoritmos de deep learning pueden llevar a cabo el auto aprendizaje de patrones ocultos dentro de los datos para hacer predicciones.

En esencia, puede pensar en deep learning como una rama de machine learning que se entrena con grandes cantidades de datos y trata con muchas unidades computacionales que trabajan en conjunto para realizar predicciones [33].

2.2.10 Elementos clave del machine learning:

Existen decenas de miles de algoritmos de aprendizaje automático y cada año se desarrollan cientos de nuevos algoritmos.

Todo algoritmo de aprendizaje automático tiene tres componentes:

- Representación: cómo representar el conocimiento. Algunos ejemplos son los árboles de decisión, los conjuntos de reglas, las instancias, los modelos gráficos, las redes neuronales, las máquinas de vectores de soporte, los conjuntos de modelos y otros.
- Evaluación: la forma de evaluar los programas candidatos (hipótesis). Los ejemplos incluyen la precisión, la predicción y el recuerdo, el error cuadrado, la probabilidad, la probabilidad posterior, el coste, el margen, la divergencia de entropía k-L y otros.
- Optimización: la forma en que se generan los programas candidatos, conocida como proceso de búsqueda. Por ejemplo, optimización combinatoria, optimización convexa, optimización restringida [27].

2.2.11 Aprendizaje inductivo:

La mayoría de las aplicaciones prácticas de la inteligencia artificial están basadas en la construcción de un modelo del conocimiento utilizado por un experto humano. Este enfoque, que ha tenido un impacto importante a partir de los ochenta, se muestra en numerosos casos de estudio (Feigenbaum et al., 1988). En la mayor parte de éstos, la tarea que ejecuta el experto es de clasificación; por ejemplo, Feigenbaum et al. (1988) citan un sistema, desarrollado por American Express, para asistir a las personas que autorizan créditos.

Existen dos formas muy diferentes de construir los modelos del conocimiento; por una parte, estos se pueden obtener entrevistando al experto o expertos apropiados. La mayoría de los sistemas basados en el conocimiento han sido elaborados de esta forma, a pesar de las dificultades relacionadas con este enfoque (Michie, 1987, 1989). Alternativamente, se pueden examinar las clasificaciones realizadas en el pasado y

construir un modelo mediante un algoritmo de aprendizaje automático que permita determinar la solución de un nuevo problema [28].

3. Procedimiento

En esta sección se describen las etapas ejecutadas para el desarrollo de esta investigación. Se hace revisión de manera general las etapas por las cuales se manipula la información para llegar al producto final, y además, la investigación previa al trabajo efectuado.

El trabajo efectuado consiste en la construcción de imágenes de energía utilizando las siluetas extraídas de los fotogramas que componen un video de una persona caminando.

A continuación, tenemos una descripción más gráfica y general de cómo funciona este proceso.

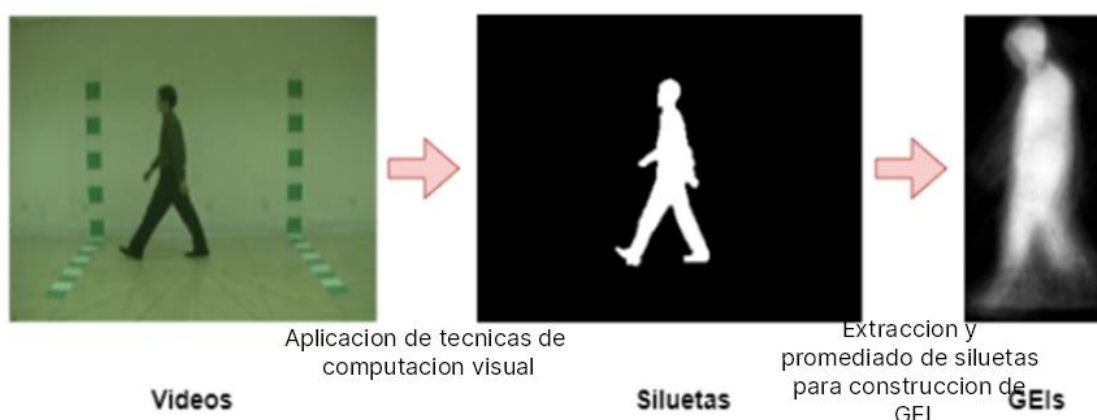


Figura 3 Creación de una imagen de energía

3.1 Investigación y practica

- **3.1.1 Trabajos previos:** Previo a la codificación, se realizó una investigación para consultar los trabajos previos en identificación biométrica y del andar humano.
- **3.1.2 Talleres de practica:** Fueron realizados pequeños talleres de practica en temas de aprendizaje maquina y de aprendizaje profundo.
- **3.1.3 Computación visual y manipulación de imágenes:** Se realizó una investigación en las maneras en las que es posible descomponer la información de un video en imágenes y modificarse la información presente en una imagen, como

es la inversión de colores, el cambiar el tamaño, la calidad y más importante la umbralización de imágenes.

Esta investigación vino en forma de un curso especializado de la paquetería OpenCV en Python, dicho curso ofrece una gran variedad de lecturas sobre herramientas para utilizar, dentro de Python, en lo que el manejo de imágenes, videos y otros componentes de computación visual se refiere.

- **3.1.4 Siluetas, imágenes de energía y operaciones morfológicas:** Se investigo y recolecto distintos métodos de cómo utilizar las siluetas obtenidas para la creación de las imágenes de energía y los métodos para aplicar operaciones morfológicas en los fotogramas (que son el resultado del descomponer un video en imágenes), esto para conseguir que la información que es de relevancia en la imagen sea resaltada para su futura manipulación.

3.2 Experimentación y datos

- **3.2.1 Datos a utilizar:** A fin de agilizar el trabajo de nuestra investigación, se decidió por la búsqueda de una base de datos o conjuntos de datos para hacer la experimentación de este trabajo, en este caso se buscaba una base de datos que contuviera un conjunto de videos con personas especialmente caminando, con una distinción de ángulos en las tomas y distintos factores en ellas, como que alguno cargara una bolsa o mochila. Al final, se obtuvo el acceso a la base de datos CASIA-B. Para lograr esto último, fue necesario hacer contacto con el investigador “Chunfeng Song”, el cual otorgo acceso a un servidor privado para la obtención de dicho set de datos.



Figura 4 Video del andar humano de la base de datos CasiaB

- **3.2.2 Prueba y validación de algoritmos:** La primera etapa de este proyecto nos suministró de las herramientas necesarias para la extracción de siluetas y generación de imágenes de energía. Sin embargo, entre las distintas técnicas que existen cada una tenía distintos parámetros y resultados, se hizo a manera general un proceso de prueba y error entre estos métodos, la primera prueba que se realizó fue con el método de binarización de Otsu.

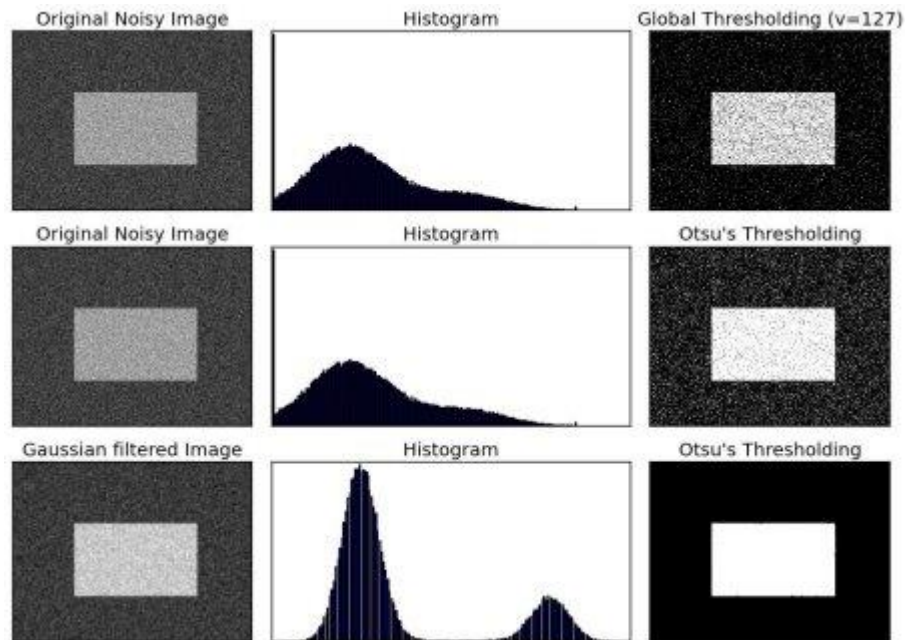


Figura 5 Uso del método de Otsu

- Sin embargo, fue obtenido como resultado que el método “Gaussian Mixture-based Background/Foreground Segmentation” [30] era el más adecuado para el objetivo de la investigación.

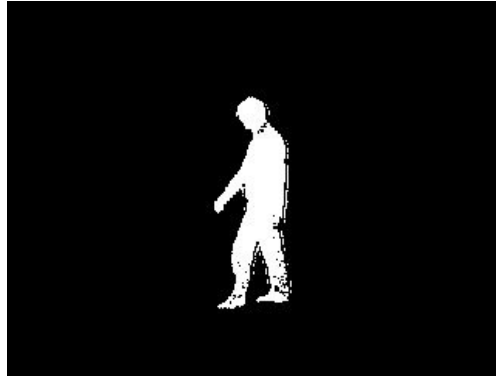


Figura 6 Imagen con silueta resaltada utilizando un tipo de Gauss



Figura 7 Otro tipo de imagen resultante del método de Gauss para la de extracción de fondos

- **3.2.3 Diseño de entrada y salida:** Al hacer uso de la base de datos CASIAB, el extractor necesitaba seguir ciertos requerimientos para poder manipular la información contenida en la base de datos, el primer paso fue diseñar nuestra entrada en base a la estructura de carpetas de CASIAB, además, tuvimos que agregar un módulo previo para tomar los videos y descomponerlos en fotogramas.

3.3 Manipulación de datos y extracción de siluetas

El funcionamiento de este trabajo esta dividido en distintos módulos, para poder lograr el producto final, es necesario tratar los datos progresivamente bajo ciertos métodos de manipulación de imágenes, estos procedimientos son van desde descomponer un video en fotogramas hasta la transformación de las matrices que construyen la imagen para obtener la información que construye las imágenes de energía.

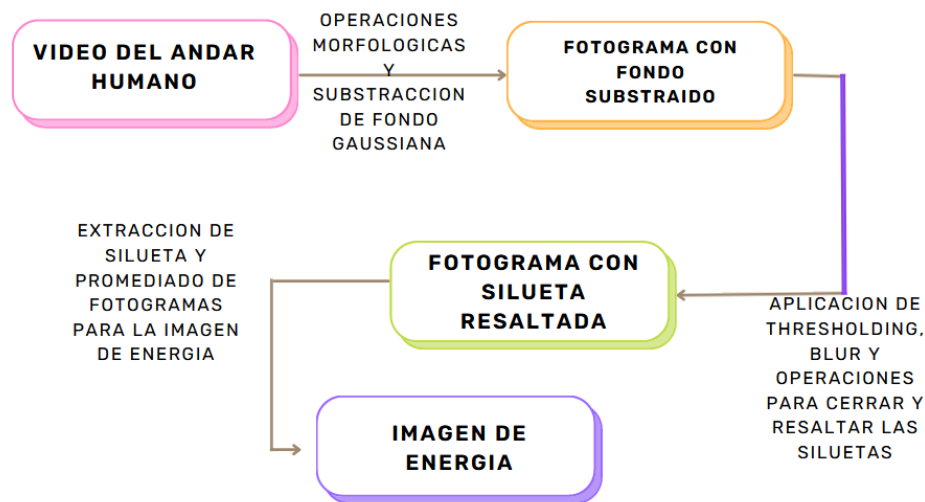


Figura 8 Funcionamiento general

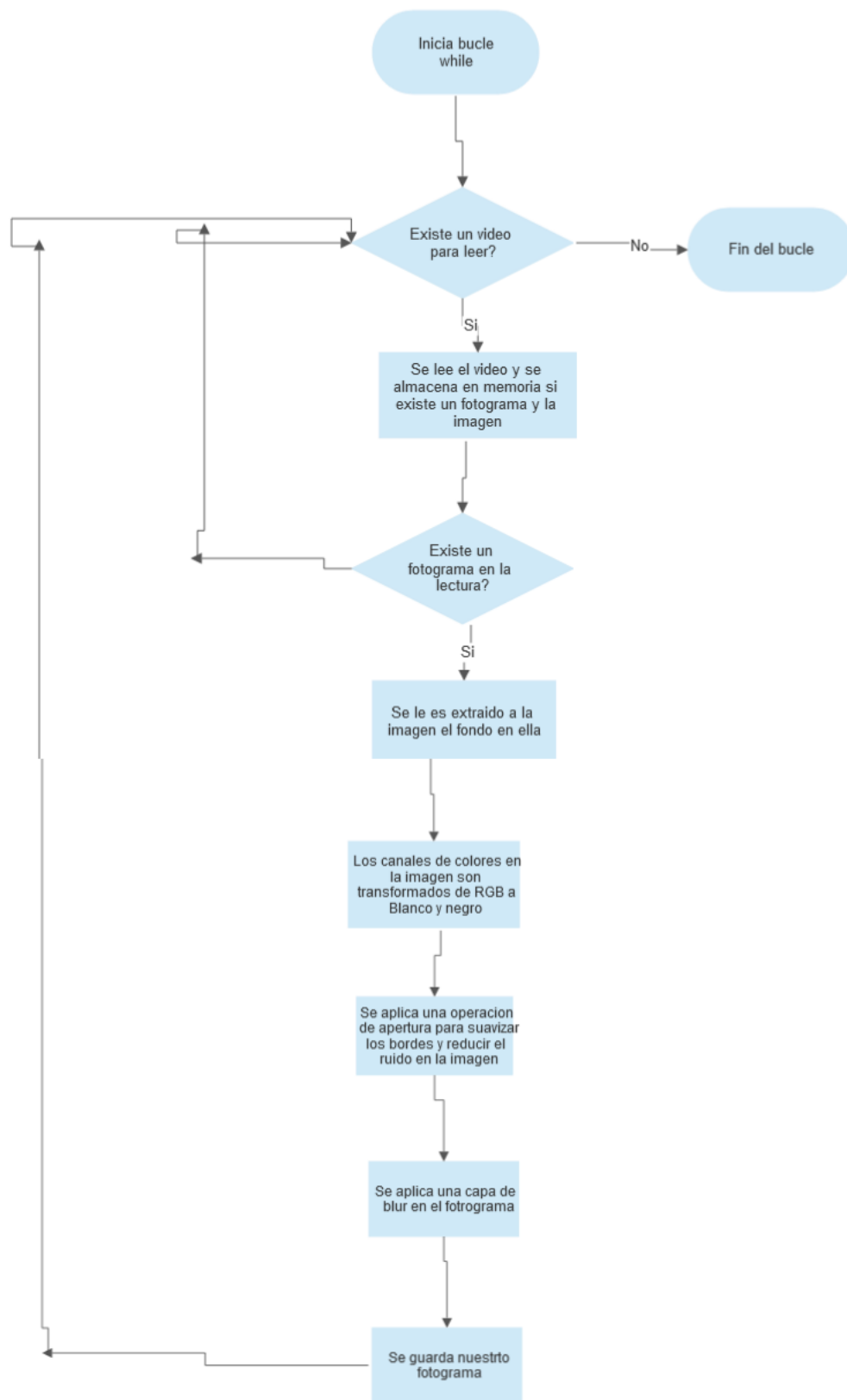
- **3.3.1 Descomposición de videos:** Se realizó una carga en memoria de los videos en donde se toma un solo fotograma de estos para su futura manipulación.
- **3.3.2 Transformación de videos por medio de operaciones morfológicas y gaussianas:** Una vez tomado el fotograma, se inicia por crear una máscara de primer plano de nuestra imagen [31], este proceso nos retorna una nueva matriz de imagen en donde el fondo se encuentra removido, después se utiliza otra

operación morfológica de apertura, esta busca por medio de la erosión de la matriz remover el ruido dentro de la imagen [32], después con una capa de desenfoque y otra capa de afinación en los bordes de nuestra imagen resultante, en este punto el ultimo filtro que se realiza en la imagen para la creación de nuestra silueta es la aplicación de una capa de thresholding, dándonos como resultado una imagen en blanco y negro, en donde toda la información innecesaria esta desechada y solo la silueta esta resaltada.



Figura 9 Silueta resultante

A continuación, el algoritmo 1 es el encargado de realizar toda la manipulación y cambios en el fotograma:



Algoritmo 1 Extracción de fondos y resaltado de silueta

3.4 Imágenes de energía

- **3.4.1 Extracción de silueta de la imagen:** En la etapa anterior, pudimos resaltar la silueta contenida en un fotograma, sin embargo, para la construcción de una imagen de energía es necesario tomar únicamente la silueta de la imagen, para lograr esta tarea se emplean 2 algoritmos, uno depende del otro, el primero se encarga de tomar los puntos de contorno de aquello que este en el centro de la imagen y/o que contenga un borde(En las imágenes con las siluetas resultantes podemos observar que solo las siluetas contienen información y el resto de la imagen carece de información alguna).

En términos generales, estos algoritmos solo buscan puntos en nuestra imagen, debido a que nos encargamos de que solo las siluetas estén presentes en estos fotogramas, solo buscan los contornos y van extrayendo la información en estos contornos.

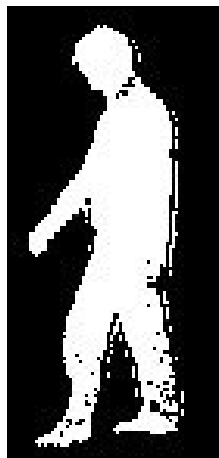


Figura 10 Silueta obtenida

Estas siluetas puras son guardadas en una lista de siluetas que corresponden a todas las siluetas que están contenidas en un video, una vez que se extraen todas las siluetas disponibles en los fotogramas que corresponden a un video es donde ocurre el promediado y la posterior construcción de imagen de energía.

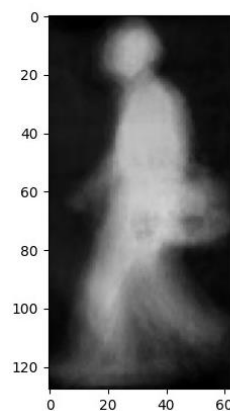
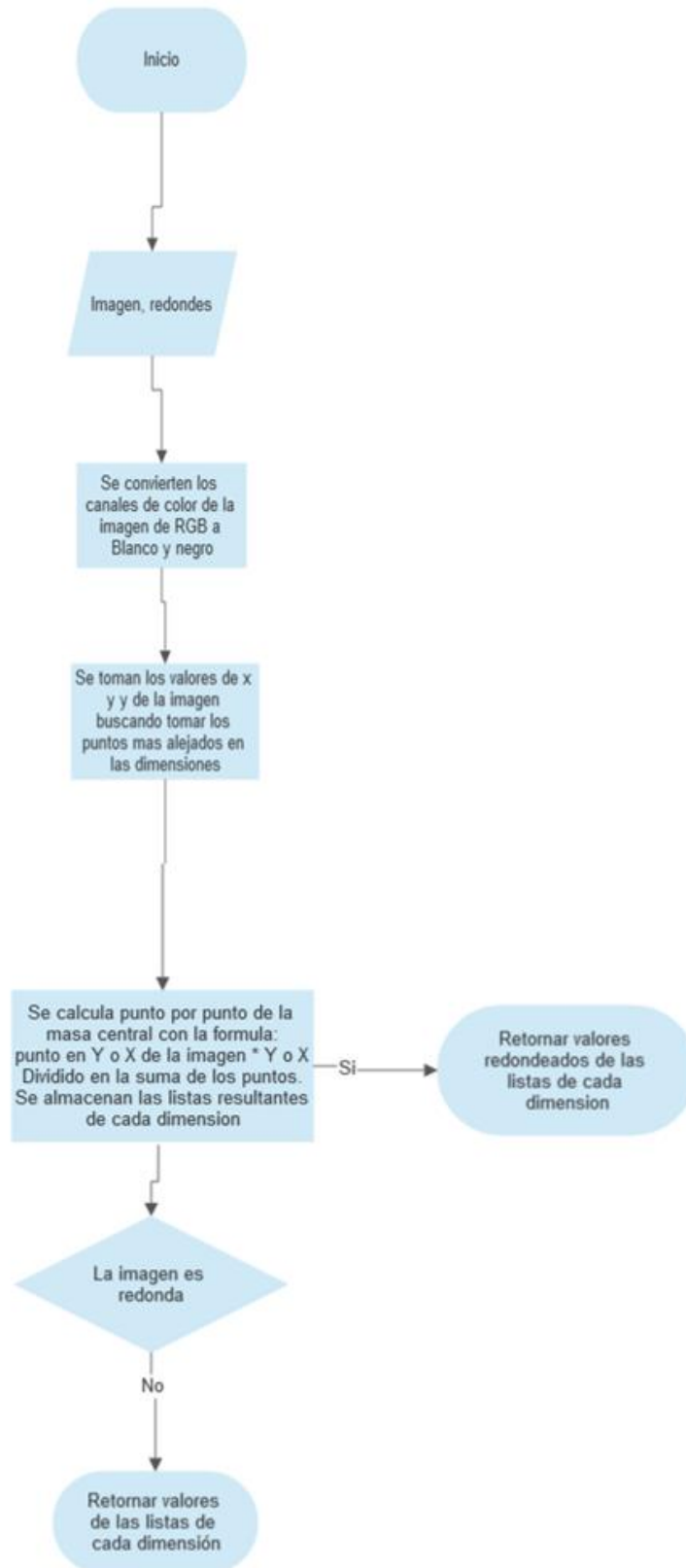
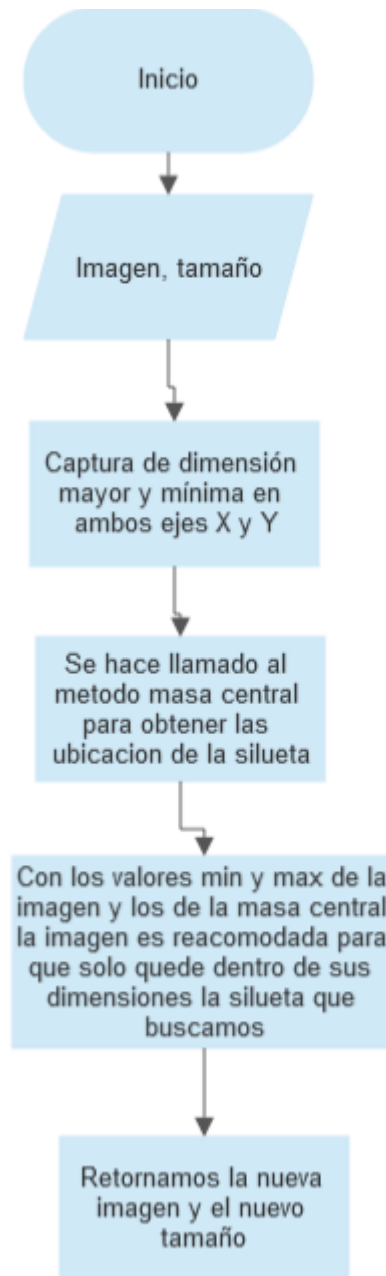


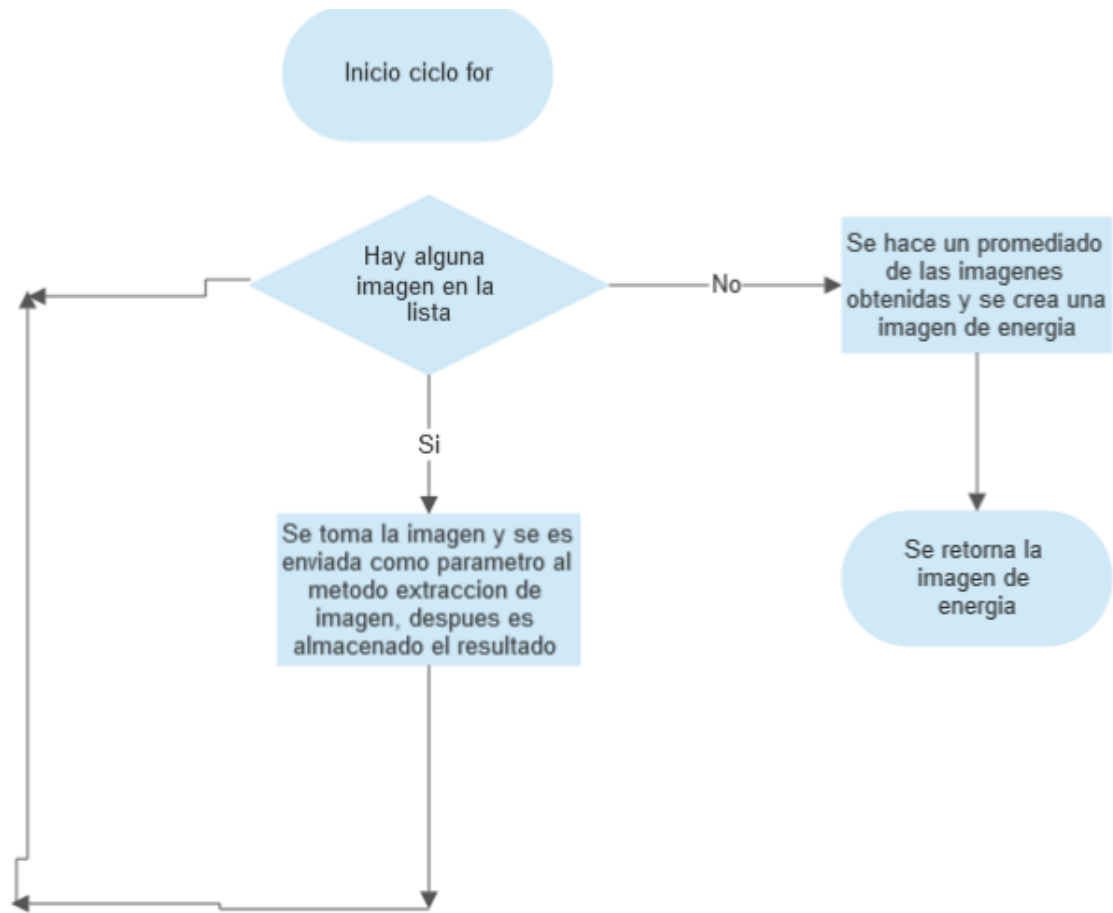
Figura 11 **Promediado de siluetas imagen de energía**



Algoritmo 2 Obtención de masa central de una imagen



Algoritmo 3 Obtención de siluetas



Algoritmo 4 Creación de imagen de energía

Los algoritmos recién descritos funcionan entre si para seguir un proceso de manipulación de imagen, a gran escala, reciben una imagen y sus dimensiones, primero se encargan de eliminar cualquier tipo de información que no sea el andar humano, una vez removida toda esta información innecesaria sigue un proceso de obtener la masa central o las dimensiones en donde se encuentre un punto que no este vacío, como se le extrajo toda la información a excepción de la silueta, se obtienen esas coordenadas y se hace una redimensión de imagen para resaltar solo la silueta contenida en el fotograma.

Por último, una vez aplicado el proceso recién descrito a cada imagen almacenada, se hace un promediado de estos fotogramas redimensionados que contienen la silueta para obtener nuestra imagen de energía de un solo

4. Resultados

Los procesos de extracción de siluetas y generación de GEI se realizaron equipo Intel Core i7 5.1 GHz con GPU Nvidia GeForce GTX 1050 TI, 16 gb de RAM Hyperx DDR4.

En promedio, cada video tiene una duración de entre 5 y 6 segundos, dando una suma de 21 horas de video. Mediante el proceso de extracción de siluetas se obtuvieron un total de 1.4 millones de siluetas de los tres grupos mencionados anteriormente.

El proceso de extracción de siluetas tuvo una duración aproximada de 32 horas, donde se obtuvo un promedio de 102 siluetas por video. Posteriormente a la extracción de la silueta, se inicia el procedimiento para generar las imágenes de energía, el cual duro 46 minutos, y se obtuvieron 13515 GEI resultantes.

Los resultados obtenidos de generación de GEI se muestran en la Tabla 1, ahí se observan los totales de GEI producidas y los porcentajes de éxito por cada uno de los grupos presentes en la base de datos CASIA-B.

Grupos	Videos	GEI	Éxito
Caminata normal	8184	8197	99.06
Vistiendo un abrigo	2728	2707	99.23
Cargando un bolso	2728	2701	99.01
Total	13640	13515	99.08

Tabla 1 Resultados obtenidos de generación de GEI por cada grupo de la base de datos CASIA-B

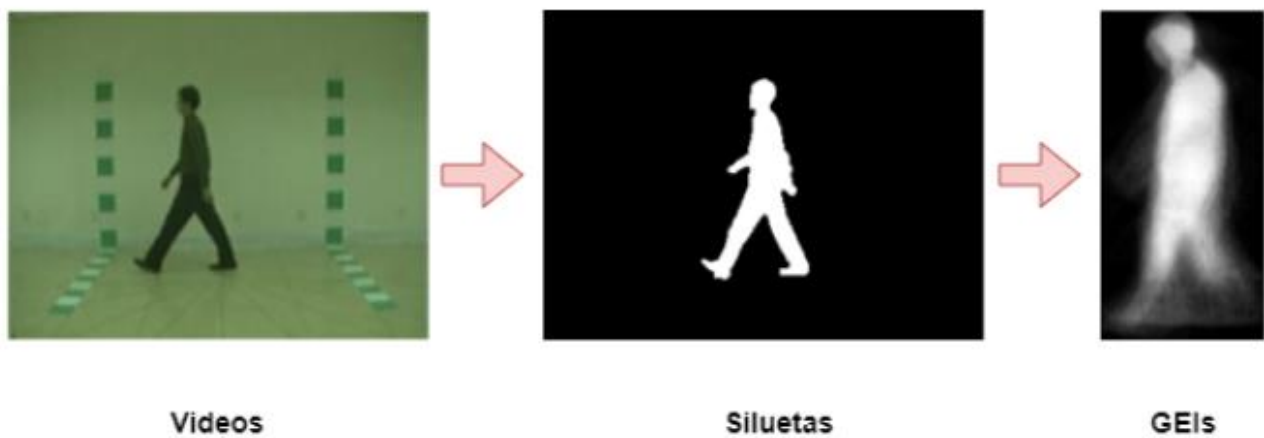


Figura 12 Preprocesamiento de información y creación de imagen de energía

Finalmente, en se muestran las GEI generadas a partir de las siluetas que se obtuvieron en el proceso de extracción de siluetas.

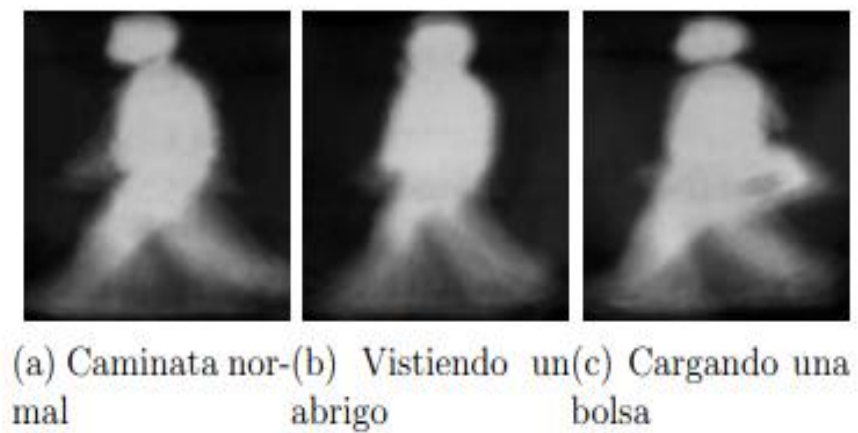


Figura 13 Tipo de imágenes de energía obtenidas

5. Conclusiones

Este trabajo propone una técnica de aprendizaje profundo, versátil y fácil de emplear para la extracción de siluetas, las CNN son una herramienta de utilización en auge y con múltiples áreas de aplicación, sobre todo en temas de tratamiento de imágenes. El identificar exitosamente a una persona por medio de sus características biométricas representa un reto que suele ser más sencillo cuando se utilizan varios enfoques biométricos al mismo tiempo, aunque la propuesta de este trabajo no representa un acercamiento directo a lo anterior dicho, el procesado de imágenes y extracción de imágenes representa la fundación para cualquiera de estos trabajos futuros.

La propuesta de este trabajo demostró una productividad en el 99 % de los videos sometidos, i.e., solo del 1 % de los videos no se logró obtener una GEI como resultado final. Esto debido a que la CNN encargada de extraer las siluetas no las detecto o no se lograron obtener los ciclos de caminata para de esa forma generar las GEI correspondiente a esta pequeña parte de los videos fallidos. Como resultado se obtuvo una base de datos de 13515 GEI y más de 1.4 millones imágenes de siluetas extraídas de los videos. Todas estas imágenes se pueden utilizar con diferentes propósitos, e.g., las siluetas servirían para proponer diferentes métodos para generación de GEI o bien, otro tipo de técnicas que requieran extracciones del ciclo de caminata de una persona; también, las GEI generadas sirven para probar propuestas de clasificadores de imágenes basados en aprendizaje profundo u otras técnicas computacionales.

6. Competencias desarrolladas

1. Implementa aplicaciones computacionales para solucionar problemas de diversos contextos, integrando diferentes tecnologías, plataformas o dispositivos.
2. Diseña, desarrolla y aplica modelos computacionales para solucionar problemas, mediante la selección y uso de herramientas matemáticas.
3. Diseña e implementa interfaces para la automatización de sistemas de hardware y desarrollo del software asociado.
4. Evalúa tecnologías de hardware para soportar aplicaciones de manera efectiva.

7. Fuentes de información

1. (2021, 3 febrero). Computación visual. Computacion Visual. Recuperado 28 de julio de 2022, de <https://techinfo.wiki/computacion-visual/>
2. Umbralización (Thresholding). (s. f.). Umbralización Thresholding. Recuperado 28 de julio de 2022, de https://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_03_04/sonificacion/cabroa_archivos/umbralizacion.html
3. Roca, J. (2022, 22 junio). Tu gráfica NVIDIA tiene cientos de ellos, pero ¿qué son los núcleos CUDA? HardZone que es nvidia cuda. Recuperado 28 de julio de 2022, de <https://hardzone.es/marcas/nvidia/nucleos-cuda/>
4. ¿Qué es Python? (s. f.). Oracle Developer. Recuperado 28 de julio de 2022, de <https://developer.oracle.com/es/python/what-is-python/>
5. Shapiro, L.G., & Costa, M.S. (1994). Appearance-Based 3D Object Recognition. Object Representation in Computer Vision.
6. Bouchrika, I. A survey of using biometrics for smart visual surveillance: Gait recognition. Surveillance In Action. (2018)
7. Han, J. & Bhanu, B. Individual recognition using gait energy image. IEEE Transactions On Pattern Analysis And Machine Intelligence. 28, 316-322 (2005)
8. Lam, T. & Lee, R. A new representation for human gait recognition: Motion silhouettes image (MSI). International Conference On Biometrics. pp. 612-618 (2006)
9. Bashir, K., Xiang, T. & Gong, S. Gait recognition using gait entropy image. (IET,2009)
10. Rodríguez, H. (2021, 28 abril). ¿Qué es OpenCV?: ¡Descubre todo acerca de la visión artificial! <https://www.crehana.com/>. Recuperado 28 de julio de 2022, de <https://www.crehana.com/blog/desarrollo-web/que-es-opencv/>
11. Hu, M., Wang, Y., Zhang, Z., Zhang, D. & Little, J. Incremental learning for video-based gait recognition with LBP flow. IEEE Transactions On Cybernetics. 43, 77-89 (2012)
12. Cunado, D., Nixon, M. & Carter, J. Automatic extraction and description of human gait models for recognition purposes. Computer Vision And Image Understanding. (2003)
13. Yam, C., Nixon, M. & Carter, J. Automated person recognition by walking and running via model-based approaches. Pattern Recognition. (2004)

14. Tang, J., Luo, J., Tjahjadi, T. & Guo, F. Robust arbitrary-view gait recognition based on 3d partial similarity matching. *IEEE Transactions On Image Processing*. (2016)
15. Lopez-Fernandez, D., Madrid-Cuevas, F., Carmona-Poyato, A., Muñoz-Salinas, R. & Medina-Carnicer, R. A new approach for multi-view gait recognition on unconstrained paths. *Journal Of Visual Communication And Image Representation*. (2016)
16. Kastaniotis, D., Theodorakopoulos, I., Economou, G. & Fotopoulos, S. Gait based recognition via fusing information from Euclidean and Riemannian manifolds. *Pattern Recognition Letters*. (2016)
17. Zhao, G., Liu, G., Li, H. & Pietikainen, M. 3D gait recognition using multiple cameras. *7th International Conference On Automatic Face And Gesture Recognition (FGR06)*. pp. 529-534 (2006)
18. Apd, R. (2021, 3 septiembre). Técnicas de inteligencia artificial: ¿cuáles son y para qué se usan? APD España. Recuperado 28 de julio de 2022, de <https://www.apd.es/tecnicas-de-la-inteligencia-artificial-cuales-son-y-para-que-se-utilizan/>
19. IBM Cloud Education. (2021, 14 septiembre). Redes neuronales. ¿Qué son las redes neuronales? Recuperado 28 de julio de 2022, de <https://www.ibm.com/mx-es/cloud/learn/neural-networks>
20. Juan Barrios. (2022, 19 enero). Redes Neuronales Convolucionales - Juan Barrios. Juan Barrios - Consultores estratégicos en Ciencia de Datos. Recuperado 28 de julio de 2022, de <https://www.juanbarrios.com/redes-neurales-convolucionales/>
21. Perceptrón multicapa. (s. f.). IBM Corp. 2017. Recuperado 28 de julio de 2022, de <https://www.ibm.com/docs/es/spss-statistics/25.0.0?topic=networks-multilayer-perceptron>
22. Liao, R., Yu, S., An, W. & Huang, Y. A model-based gait recognition method with body pose and human prior knowledge. *Pattern Recognition*. 98 pp. 107069 (2020)
23. Makiyara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T. & Yagi, Y. Gait Recognition Using a View Transformation Model in the Frequency Domain. *Computer Vision – ECCV 2006*. pp. 151-163 (2006)
24. Geografía, E. D. N. I. Y. (s. f.). *Incidencia delictiva*. Recuperado 25 de octubre de 2022, de <https://www.inegi.org.mx/temas/incidencia/>

25. Machine Learning. (s. f.). IBM Analytics. Recuperado 28 de julio de 2022, de <https://www.ibm.com/mx-es/analytics/machine-learning>
26. Delgado, D. O. (2019, 17 septiembre). ¿Qué es Tensorflow? OpenWebinars.net. Recuperado 28 de julio de 2022, de <https://openwebinars.net/blog/que-es-tensorflow/>
27. Operaciones morfológicas - MATLAB & Simulink - MathWorks América Latina. (s. f.). <https://la.mathworks.com/help/images/morphological-filtering.html>
28. Eliminación gaussiana. (s. f.). StudySmarter ES. <https://www.studysmarter.es/resumenes/matematicas/numeros-y-algebra/eliminacion-gaussiana/>
29. (2019, 8 febrero). Parámetros biométricos. Kimaldi Parámetros biométricos. Recuperado 28 de julio de 2022, de https://www.kimaldi.com/blog/biometria/parametros_biometricos/ [29]
30. ¿Qué es el aprendizaje no supervisado? (s. f.). TIBCO Software. Recuperado 28 de julio de 2022, de <https://www.tibco.com/es/reference-center/what-is-unsupervised-learning>
31. Aprendizaje Supervisado. (s. f.). MATLAB & Simulink. Recuperado 28 de julio de 2022, de <https://es.mathworks.com/discovery/supervised-learning.html>
32. ¿Qué es la inteligencia artificial? | Microsoft Azure. (s. f.). Microsoft Azure. <https://azure.microsoft.com/es-es/resources/cloud-computing-dictionary/what-is-artificial-intelligence/>
33. Education, I. C. (2022, 30 marzo). Deep Learning. <https://www.ibm.com/cloud/learn/deep-learning>