

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA



**Trabajo de investigación acerca de sistemas de grabación automática para la
detección, seguimiento y conteo de peatones en intersecciones semaforizadas**

**TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO
ACADÉMICO DE BACHILLER EN CIENCIAS CON MENCIÓN EN
INGENIERÍA MECATRÓNICA**

AUTOR

Martín Giancarlo Leiva Ottone

ASESOR

Pedro Moises Crisostomo Romero

Julio, 2020

Resumen

La información acerca del flujo peatonal es necesaria como métrica para dirigir diversos programas relacionados al transporte no motorizado (Lindsey, Nordback, & Figliozi, 2016); sin embargo, actualmente la manera más común de obtener esta información es a través del conteo manual, el cual cuenta con diversas desventajas como ser una medida de corto plazo y de costo relativamente elevado (U.S Department of Transportation, 2016). Ante tales circunstancias, el objetivo general del presente trabajo académico es encontrar un concepto de solución que nos permita automatizar el proceso de conteo peatonal. Esta automatización se logra mediante un sistema de grabación que cuente con una cámara móvil que le permita alternar entre dos cruces peatonales adyacentes dependiendo del color del semáforo.

Como parte de la elaboración conceptual de la propuesta, se realiza una investigación del estado del arte de temas relacionados a la detección, seguimiento y conteo de peatones, así como la automatización de plataformas para cámaras. Esta información es utilizada para poder definir los requerimientos de dicho sistema, sus principales funciones y proponer alternativas de solución. Finalmente, utilizando un análisis económico-técnico se elige el concepto de solución óptimo.

Índice de contenido

Índice de figuras	iii
Índice de tablas	iv
Introducción.....	1
Capítulo 1 Marco Problemático.....	3
1.1 Problemática	3
1.1.1 Razones de la falta de información peatonal	3
1.1.2 Problemática principal por tratar	4
1.1.3 Razones sociales: Preferencia del vehículo sobre el peatón	5
1.2 Propuesta de solución	6
1.2.1 Alcance	6
1.2.2 Objetivos.....	7
1.3 Grupos de interés	7
1.4 Metodología.....	9
Capítulo 2 Estado del arte.....	11
2.1 Detección de peatones	11
2.1.1 Detectores de dos etapas.....	11
2.1.2 Detectores de una etapa	17
2.2 Seguimiento de peatones	20
2.3 Conteo de peatones.....	23
2.4 Plataformas móviles para cámaras.....	30
2.5 Control de plataformas móviles.....	33
Capítulo 3 Diseño de la propuesta.....	39
3.1 Lista de requerimientos.....	40
3.2 Caja negra del sistema	41
3.2.1 Función principal.....	41
3.2.2 Entradas y salidas	42
3.3 Estructura de funciones.....	43
3.3.1 Subsistemas	45
3.4 Matriz Morfológica.....	49
3.5 Conceptos de solución	54
3.5.1 Concepto de solución 1.....	54
3.5.2 Concepto de solución 2.....	55
3.6 Evaluación de los conceptos de solución.....	59
3.6.1 Criterios Técnicos.....	60
3.6.2 Criterios Económicos.....	61
3.6.3 Obtención del concepto óptimo	61
Conclusiones.....	64
Bibliografía.....	65

Índice de figuras

Figura 1-1: Norma VDI 2221.Adaptado	10
Figura 2-1: Características tipo Haar	14
Figura 2-2a, 2-2b, 2-2c, 2-2d, 2-2e y 2-2f : Resultado del método 1, 2 , 3, 4, 7 y 9 respectivamente.....	17
Figura 2-3: Ejecución del modelo YOLO.....	18
Figura 2-4: Funcionamiento del seguimiento usando filtro de partículas.....	21
Figura 2-5a: Implementación del filtro Kalman y del filtro de partículas en casos sin oclusión.	22
Figura 2-6 Implementación de filtro Kalman y algoritmo Camshift	23
Figura 2-7: Ejemplo de detección de los puntos de interés	27
Figura 2-8: Línea principal y línea auxiliar	28
Figura 2-9 Matrices espaciotemporales: izquierda: línea principal, derecha: línea auxiliar ..	28
Figura 2-10 Resultados del análisis	30
Figura 2-11 Ejemplos de sistemas Pan Tilt y PTZ de izquierda a derecha.....	31
Figura 2-12a y Figura 2-12b: Ejemplo de articulaciones Print and fold y Plataforma de cámara impresa en 3D	32
Figura 2-13 Brazo robótico, cámaras PTZ con sus respectivos sistemas de coordenadas	33
Figura 2-14: Demostración del control por seguimiento.	34
Figura 2-15: Resultados del control de acuerdo con un objeto móvil	35
Figura 2-16 Diagrama de bloques del sistema de reconocimiento facial implementado	36
Figura 2-17: Respuesta de la plataforma a una cara en movimiento	37
Figura 2-18: Respuesta en lazo cerrado del eje horizontal y vertical respectivamente	37
Figura 3-1,3.1b y 3.1c: Esquema del posicionamiento del sistema, perspectivas del sistema con respecto a los cruces peatonales	39
Figura 3-2: Caja negra del sistema.....	41
Figura 3-3: Estructura de funciones.....	44
Figura 3-4 Concepto de solución 1	54
Figura 3-5 Concepto de solución 2	57
Figura 3-6 Concepto de solución 3	58
Figura 3-7 Gráfico técnico-económico entre las soluciones.....	63

Índice de tablas

Tabla 2-1: Comparación de métodos de detección de peatones (Rauf, Shahid, Ziauddin, & Safi, 2017)	16
Tabla 2-2: Comparación entre detectores de puntos de interés (Jeong & Choi, A Comparison of Keypoint Detectors in the Context, 2016) (Adaptado).....	26
Tabla 2-3: Resultados del proceso (Yu, Gong, Yang, & Li, 2014)	29
Tabla 3-1 Lista de requerimientos. Elaboración propia.....	40
Tabla 3-2 Matriz Morfológica. Elaboración propia.....	50
Tabla 3-3 Pesos relativos. Elaboración propia.....	59
Tabla 3-4. Escala de efectividad de las soluciones. Elaboración propia	59
Tabla 3-5 Criterios Técnicos. Elaboración propia	60
Tabla 3-6 . Criterios Económicos. Elaboración Propia.....	61
Tabla 3-7. Evaluación técnica-económica. Elaboración propia.....	62

Introducción

Las tendencias mundiales las últimas 2 décadas han revelado que tanto investigadores como diferentes gobiernos se han interesado en mejorar las condiciones del transporte peatonal (Lindsey, Nordback, & Figliozi, 2016), en cierta manera debido a que actualmente el transporte está experimentando un cambio de un enfoque de uno basado en los vehículos, a otro en el cual los transeúntes son la prioridad (Orellana, Hermida, & Osorio, 2017). Ante estas circunstancias diferentes autoridades han decidido invertir en infraestructura urbana como cruces peatonales, semáforos entre otros. Igualmente, las ciudades interesadas en promover la movilidad urbana inician elaborando un Plan de Movilidad el cual ayuda a planificar el transporte estableciendo estrategias y regulaciones para mejorar el transporte de los ciudadanos (Fernández-Garza & Hernández-Vega, 2019).

Tanto para el planeamiento de dichas políticas públicas como para medir su éxito es necesario tener datos cuantitativos y detallados del flujo peatonal. Sin embargo, actualmente existe una escasa información peatonal, por ejemplo, en Perú, la encuesta “Lima como vamos”, la cual está orientada a conocer la percepción sobre la calidad de vida en Lima y Callao, ahondó por primera vez acerca de los distintos medios que las personas emplean para ir a su trabajo en la edición del 2017 (Instituto de Opinión Pública-PUCP, 2018). A nivel mundial, el caso es similar, la recolección de datos de transporte no motorizado es reciente y la experiencia en el monitoreo peatonal es limitada, en gran medida debido a que el conteo manual es costoso y laborioso. (Fernández-Garza & Hernández-Vega, 2019). En cuanto al conteo automático de tráfico no motorizado, los métodos que involucran procesamiento de imágenes son considerados tecnologías emergentes y están siendo adaptadas (aunque aún no del todo) de sus contrapartes en el monitoreo de tráfico motorizado debido a que tienen un potencial de tener alta precisión

pero aún sus algoritmos relacionados siguen madurando (U.S Department of Transportation, 2016).

Teniendo en cuenta el contexto actual descrito previamente, el presente proyecto se enfoca en plantear un concepto de solución que permita automatizar la obtención de datos acerca del flujo peatonal. Se utiliza procesamiento de imágenes en un sistema de grabación que cuente con una cámara móvil la cual se alterna automáticamente entre dos cruces peatonales dependiendo del color del semáforo. Este sistema se encuentra instalado en un poste y contempla una estructura que cuente con protección de grado IP54 la cual le permite protegerse de polvo y lluvia. La alimentación es a través de baterías y se activará con un botón de encendido. Los datos de entrada del sistema, como el inicio del proceso, su duración, la ubicación aproximada de los cruces peatonales, entre otros, son ingresados por el operario mediante una interfaz remota. Una vez ingresados estos datos el sistema comenzará a realizar el conteo de peatones hasta el final de la duración programada. Terminado el proceso, los datos acerca del flujo peatonal son enviados al operario y mostrados en la interfaz.

El presente trabajo de investigación está organizado en 3 capítulos. En el capítulo 1 detallará información acerca de la problemática asociada a la automatización del conteo peatonal. Seguidamente, se señalarán el objetivo general y los objetivos específicos del presente documento académico. Asimismo, se explicará la metodología a utilizar y mencionarán los grupos de interés. En el capítulo 2 se analizará el estado del arte de la tecnología relacionada. En el capítulo 3 se tratarán los temas de lista de requerimientos, estructura de funciones, matriz morfológica, conceptos de solución, así como la evaluación de estos y la elección del concepto óptimo. Por último, se detallan conclusiones para la consideración de futuros trabajos relacionados.

Capítulo 1 Marco Problemático

En el presente capítulo se examinan los aspectos de la problemática relacionada al flujo peatonal, la propuesta de solución planteada, los objetivos, alcance, la identificación de los usuarios, clientes e interesados, así como especificar el plan metodológico a seguir.

1.1 Problemática

En los últimos veinte años los gobiernos de turno, en todos los niveles, se han interesado en la infraestructura para el viaje no motorizado, así como mejorar la gestión del transporte; pero en la mayoría de los casos las iniciativas han sido concretadas sin tener en cuenta datos cuantitativos del flujo peatonal debido a la falta de estos datos. (Lindsey, Nordback, & Figliozzi, 2016). La información acerca del conteo de peatones es esencial para la toma de decisiones relacionadas con el planeamiento urbano, corrección del ciclo semafórico, apropiada demarcación de vías, la mejora de la señalización, así como medir el éxito de dichas políticas (Yang, Kaan, & Bartin, 2010).

1.1.1 Razones de la falta de información peatonal

Lamentablemente, la falta de esta información se puede atribuir a la dificultad de obtenerla con alta calidad. Para conocer las razones es necesario conocer la situación del monitoreo del tráfico no motorizado, el cual es explicado en la Guía de Monitoreo del Tráfico, elaborado por el departamento de transporte de los Estados Unidos. De este manual se pueden identificar los siguientes factores.

El primero de estos es el costo, uno de los métodos más comunes para el conteo, el cual se realiza mediante observadores que ejecutan el trabajo de manera manual. Este no solo es uno de los métodos más usados, sino también el más costoso debido a la paga que deben recibir los trabajadores. Además, este método tiene la desventaja de ser poco

preciso para cuentas de larga duración, haciendo este método recomendable para recolección de datos a corto plazo.

Otra de las razones a considerar es la corta duración de las cuentas, prioritariamente debido a la percepción de dificultad con respecto al conteo automático de peatones. Esto se vuelve un problema debido a que estas cuentas de corta duración pueden introducir un error significativo cuando el tráfico peatonal es bajo e innatamente variable. Es debido a esto que es necesario el uso de cuentas más largas para poder establecer patrones y las bases estadísticas para la extrapolación de los conteos de menor duración.

Por otro lado, se señala que a diferencia que el conteo de vehículos, las tecnologías enfocadas en peatones aún están desarrollándose y los porcentajes de errores relacionados con diferentes métodos aún no son bien conocidos, haciendo difícil el desarrollo de procedimientos que manejen o reduzcan estos errores. Entre los principales retos de estas tecnologías se encuentran la poca predictibilidad de los transeúntes y la dificultad para distinguir peatones en grupos numerosos (U.S Department of Transportation, 2016).

Finalmente, es importante mencionar que la toma de datos utilizando video es generalmente estática y limitada a una perspectiva fija, lo cual dificulta monitorear objetos alejados del ángulo de la cámara. Esto se podría solucionar automatizando cámaras móviles, pero en la actualidad estas son mayormente controladas manualmente o mediante secuencias preprogramadas. (Kim, Kim, & Sungjoo, 2019).

1.1.2 Problemática principal por tratar

Teniendo en cuenta los factores explicados anteriormente, se ha decidido que el presente trabajo de investigación aborde la problemática de ***Falta de automatización para la toma de datos de flujo peatonal.*** Esto se debe a que este enlazará los problemas de costo, corta duración y falta de tecnología.

1.1.3 Razones sociales: Preferencia del vehículo sobre el peatón

Teniendo ya en cuenta los detalles técnicos de la problemática, es necesario abordar el aspecto social para llegar a un mayor entendimiento de la situación. A manera de ejemplo, se volverá a mencionar la Guía de Monitoreo del Tráfico, en especial la sección de monitoreo de tráfico no motorizado, debido a que su contenido hace evidente que en algún momento existió la preferencia del vehículo sobre el peatón. Esto se puede observar cuando se menciona que esta edición del manual (la cual fue publicada en el año 2016) es la primera en incluir información acerca del monitoreo de flujo peatonal debido a que esta no es tan sistemática, extensa, ni comprensible como la del monitoreo de flujo vehicular, aun cuando el movimiento peatonal le precede al motorizado (U.S Department of Transportation, 2016).

Esta preferencia que se le tiene al vehículo sobre el peatón se puede explicar entendiendo que en estos últimos años nos encontramos en un periodo de transición de un paradigma del transporte, el cual favorece la planificación enfocada en vehículos, al paradigma de la movilidad, que considera una planificación que tiene como base las necesidades de acceso del ser humano. Este paradigma de la movilidad ha cobrado fuerza, debido a factores como el gran flujo de datos sobre geolocalización que existen en la actualidad, el cual no se podía imaginar en épocas pasadas, y por los nuevos modelos de análisis de vehículos y peatones, los cuales utilizan fuentes de datos como las provenientes de video capturado por cámaras de seguridad, siendo estos factores utilizados para la mitigación de impactos negativos producidos por años de planificación solo en función al vehículo privado (Orellana, Hermida, & Osorio, 2017).

1.2 Propuesta de solución

Se plantea una investigación enfocada en sistemas de grabación automática para la detección, seguimiento y conteo de peatones en intersecciones semaforizadas con el fin de proponer un concepto de solución óptimo a la problemática descrita en la sección 1.1

El concepto de solución óptimo debe ayudar en la automatización de la toma de datos de flujo peatonal, lo cual facilita la toma de decisiones, las cuales pueden ser asistidas por métodos de análisis y/o predicción de tráfico peatonal, estos beneficiándose de los datos recopilados (Tettamanti, Csikòs, Kis, Viharos, & Varga, 2017). Entre las posibles decisiones a tomar, se puede tener en cuenta optimizar los ciclos semafóricos, evaluar la construcción de intersecciones diagonales como ya se ha visto en el distrito de Miraflores, teniendo como ejemplo el cruce de la avenida José Larco con la avenida Ernesto Diez Canseco (Peña, 2019), o verificar la efectividad de dichas medidas.

1.2.1 Alcance

- El concepto de solución óptimo involucra un sistema de grabación automática que cuente con la capacidad de direccionarse automáticamente.
- El concepto de solución monitorea dos cruces peatonales simultáneamente durante el día.
- El sistema para conceptualizar se direcciona al cruce peatonal respectivo mediante un sistema de control cuyos datos de entrada estarán relacionados con el video grabado, el estado del semáforo y el comportamiento peatonal.
- La investigación realizada contempla la exploración del estado del arte de tecnologías relacionadas la detección, seguimiento y conteo de peatones entre otros temas relacionados.

1.2.2 Objetivos

En esta sección se presentan el objetivo general y los objetivos específicos del presente documento académico.

Objetivo General:

- Proponer un concepto de solución óptimo para la problemática de la falta de automatización en la toma de datos de flujo peatonal.

Objetivos Específicos:

- Identificar e investigar acerca de las tecnologías para la detección, seguimiento, conteo de peatones y control del sistema.
- Reconocer los requerimientos necesarios del sistema a conceptualizar.
- Identificar las funciones específicas de procesamiento de imágenes, control y movimiento que el concepto de solución debe realizar.
- Proponer diversos conceptos de solución con el fin de encontrar el óptimo.
- Orientar el desarrollo del proyecto hacia los objetivos del desarrollo sostenible, especialmente el número 11, el cual plantea ciudades y comunidades sostenibles, solucionando problemas como el de la congestión.

1.3 Grupos de interés

En la presente sección se identifican a los potenciales clientes, usuarios, y otros interesados, así como sus necesidades.

Para poder establecer los potenciales clientes y usuarios de este producto, es necesario resaltar que la finalidad los datos obtenidos por los métodos explorados en esta investigación, es la toma de decisiones relacionadas al transporte peatonal y vehicular, así como la verificación de los resultados de estas. Por esta razón, los principales clientes

a considerar son las municipalidades que necesiten los datos necesarios para cumplir con los propósitos mencionados previamente, siendo los usuarios los empleados de las oficinas municipales de tráfico encargados de la instalación y monitoreo de estos sistemas, así como los responsables de analizar la información que este sistema les proporciona. Entre los otros interesados se pueden considerar a las empresas, debido a que este sistema busca proporcionar información relevante para la disminución de la congestión y tiempos de viaje, los cuales según estudios disminuyen la productividad (Boarnet, 1997; Hymel, 2009; Zheng, Capra, Wolfson & Yang, 2014). La principal necesidad de estos interesados es que la información brindada logre aliviar la congestión que se presenta.

En el caso de la ciudad de Lima, las posibles municipalidades a considerar son las de Miraflores y San Isidro, debido a que estas forman el eje con mayor número viajes metropolitanos. En el caso de Isidro, este distrito recibe el doble de su población en viajes. Además, aunque este presenta semáforos sincronizados, no ha realizado estudios para olas verdes (Municipalidad de San Isidro, 2016), las cuales permiten un continuo flujo del tráfico por el medio de coordinación de semáforos. En el caso de la Municipalidad de Miraflores, esta cuenta con cuatro cruces en diagonal ubicados en las intersecciones de la avenida José Larco con la calle Schell, la avenida José Larco con la avenida Alfredo Benavides, la avenida José Larco con la avenida Ernesto Diez Canseco y la calle Schell con el pasaje Los Pinos, los cuales fueron implementados en Julio del 2019 (Peña, 2019), y se podría necesitar datos cuantitativos para comprobar su efectividad. También cabe resaltar que en su plan de movilidad urbana del 2017 se menciona que recibe 3.5 viajes de visitantes por cada habitante. Otro punto importante que destacar es el hecho que uno de los objetivos específicos de este plan es “Cambiar los métodos de análisis y los procedimientos de intervención reenfocándolos hacia la movilización de personas y no

de vehículos” (Municipalidad de Miraflores, 2017), siendo este proyecto una potencial herramienta para dichos métodos de análisis requeridos por la municipalidad.

1.4 Metodología

Para el presente proyecto se adaptará la metodología de diseño de sistemas técnicos y productos (VDI 2221), la cual se puede ver en la *Figura 1-1*. Debido al alcance de este trabajo de investigación se utilizará la norma hasta el paso número 3, en el cual se buscan conceptos de solución. La aplicación de la metodología es explicada a continuación.

- Definición de la problemática, incluye la descripción del problema, sus causas e impacto.
- Determinación del alcance y planteamiento de objetivos generales y específicos, para este caso, conocer la importancia del flujo peatonal en los estudios de tránsito.
- Búsqueda de principios de solución revisando el estado del arte asociado al sistema a desarrollar. Se considerarán algoritmos de detección, seguimiento y conteo de peatones, así como información acerca del diseño y control de plataformas móviles para cámaras, entre otros temas.
- Establecer los requerimientos para así poder dividir la función principal del sistema en módulos más pequeños y realizables.
- Determinar la estructura de funciones.
- Realizar una matriz morfológica para el diseño de los módulos del sistema.
- Proponer conceptos de solución y elegir el óptimo, teniendo en cuenta criterios técnicos y económicos.
- Redactar el documento final.

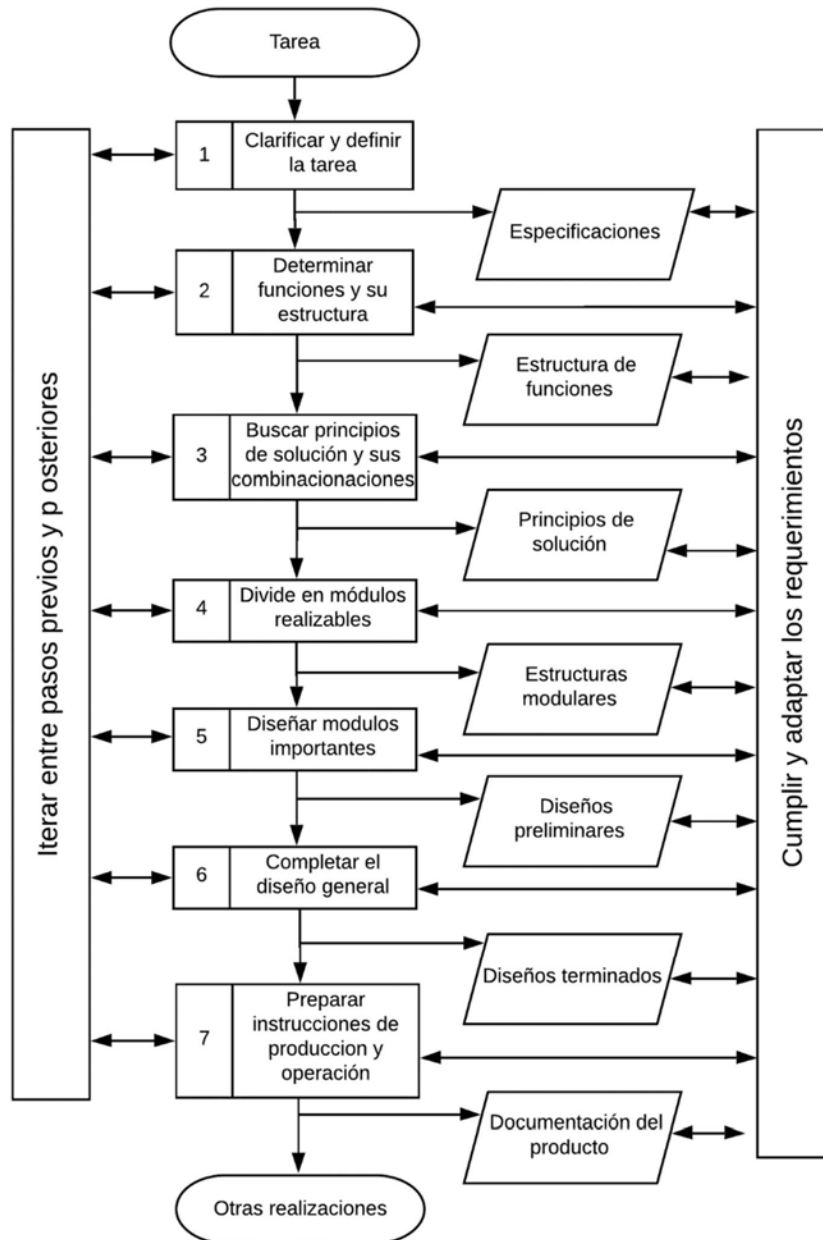


Figura 1-1: Norma VDI 2221. (Pahl, Beitz, Feldhusen, & Grote, 2007). Adaptado.

En este capítulo se ha tratado del marco problemático, el cual además de describir la problemática, se especifican los grupos de interés, los objetivos y alcances de la investigación además de explicar la metodología a utilizar. Esta información sirve para establecer el contexto en el cual se desarrolla el presente documento académico.

Capítulo 2 Estado del arte

En el presente capítulo, se muestra una recopilación de información acerca de los diversos temas relacionados con este trabajo académico con el fin de aplicar dichos conocimientos en el diseño. Los temas por tratar son los de detección, seguimiento y conteo de peatones, así como el diseño y control de plataformas móviles para cámaras.

2.1 Detección de peatones

Actualmente existen dos tipos de detectores de peatones. Por un lado, tenemos los detectores de dos etapas, como HOG (Histogram of Oriented Gradients) o Faster R-CNN (Region-based Convolutional Neural Networks), las cuales en su primera etapa extraen características de una imagen, obteniendo regiones de interés. Luego, en su segunda etapa, envían estas regiones a un clasificador, el cual determina a que clase pertenecen los objetos. Estos son los modelos con mayor precisión, pero también los más lentos. Por otro lado, existen los detectores de una etapa, como YOLO (You Only Look Once) y SSD (Single Shot Detector) los cuales tratan la detección como un problema de regresión. Estos métodos son mucho más veloces, pero tienen una menor precisión (Soviany & Tudor, 2018).

2.1.1 Detectores de dos etapas

Para entender los métodos de detección de peatones en dos etapas, se muestra una recopilación de las técnicas más utilizadas de extracción de características y clasificadores.

Técnicas de extracción de características: Estos se encargan de obtener un conjunto de características independientes entre sí, con las cuales se podrá clasificar un patrón posteriormente. Las técnicas HOG, CSS, LUV, LBP, flujo óptico y características tipo Haar, son las descritas a continuación.

- HOG: El algoritmo de Histogramas de Gradientes orientadas (HOG), es el núcleo de diversos sistemas de detección de peatones y objetos. Este método se compone en 4 pasos. Cálculo de la gradiente, almacenamiento de la orientación, bloques de descripción y bloques de normalización. En el primer bloque se detectan las gradientes y consiguientemente las esquinas. En la segunda fase se divide la imagen en celdas y se calcula el histograma de orientaciones de la gradiente en cada celda. En la tercera etapa, se juntan las celdas en bloques basándose en la orientación de su gradiente, cada bloque tiene un descriptor los cuales son usados para detectar los bounding boxes. Finalmente, se normalizan los histogramas para evitar problemas de iluminación (Fernandes, Weigel, Jung, & Navaux, 2016).
- LBP: El método conocido como Patrón Local Binario (LBP) es un extractor de características conocido por tener un desempeño excepcional como descriptor de texturas y una de sus ventajas principales es su invarianza ante el cambio de nivel de gris monótono, lo cual puede complementar al HOG extrayendo el ruido del fondo, esto se debe a que este método se basa en diferenciar patrones uniformes de no uniformes mediante la realización de histogramas de celdas individuales (Wang, Han, & Yan, 2009).
- CSS: Auto similitud de color es una técnica que consiste en obtener la semejanza en el histograma de color entre bloques de pixeles que pueden estar espacialmente separados. Esta información puede ser usada para un proceso discriminatorio posterior (Zeng, Ouyang, & Xiaogang, 2013).
- LUV: Es un espacio de color el cual puede llegar a ser más discriminativo que el canal HSV en cuanto a su uso en histogramas de gradiente. (Zhang, Bauckhage, & Cremers, 2014). Este se enfoca que la información de luminosidad, saturación y matiz (Rauf, Shahid, Ziauddin, & Safi, 2017).

- Flujo Óptico: Este método permite extraer una característica de la imagen relacionada con el movimiento de objetos, superficies y esquinas causado por el movimiento relativo entre el observador y la escena. Normalmente se utilizan fotogramas adyacentes para el procesamiento (Rauf, Shahid, Ziauddin, & Safi, 2017).
- RCNN: Las redes neuronales convolucionales basadas en regiones es un algoritmo que utiliza un grupo de recuadros para una imagen y analiza cada recuadro para saber si alguno contiene un objeto. Emplea el método de búsqueda selectiva para escoger secciones de una imagen. En cada sección se seleccionan características utilizando una red neuronal (Adarsh, Rathi, & Kumar, 2020).
- Fast RCNN: Este algoritmo emplea solo un modelo para obtener las características de diferentes regiones de la imagen. Luego distribuye las regiones en distintas categorías basadas en estas características, obteniendo recuadros de interés (Adarsh, Rathi, & Kumar, 2020).
- Faster RCNN: Faster RCNN es una variante de Fast RCNN, la diferencia entre estos dos métodos es que Faster RCNN implementa una red de regiones propuestas (RPN) en vez de usar una técnica de búsqueda selectiva para encontrar estas regiones. La RPN tiene como entrada un mapa de características de la imagen, la cual utiliza para obtener recomendaciones. Usualmente este método demora 10 veces menos que Fast RCNN (Adarsh, Rathi, & Kumar, 2020).
- Características tipo Haar: Este clasificador utiliza un conjunto de plantillas (ver *Figura 2-1*) las cuales representan las características tipo Haar de manera que se pueda extraer información de la textura de una clase de objetos. Esto se logra codificando en una imagen las diferencias de las intensidades promedio entre dos

regiones rectangulares adyacentes, capturando las similitudes entre las instancias analizadas y las plantillas (Jimenez, Martin, Uc, & Espinosa, 2017).

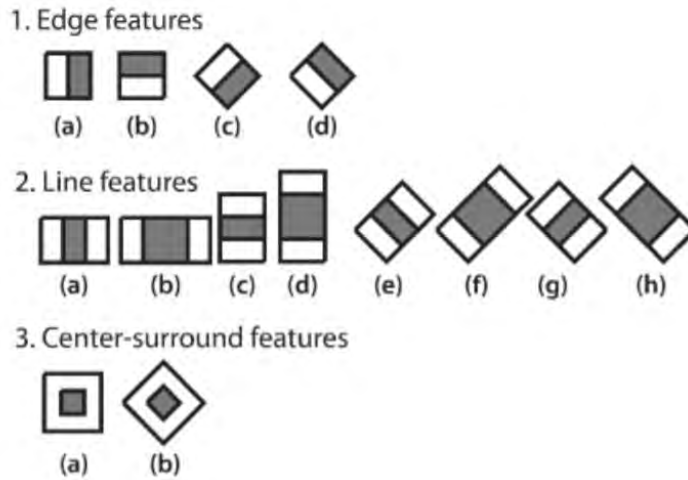


Figura 2-1: Características tipo Haar (Uddin & Akhi, 2016).

Clasificadores: Son los métodos que se utilizan para definir en qué clase se ubica cada patrón dado, debido a sus características. Los clasificadores por tratar son: SVM, Deep Learning, Árbol de decisiones y AdaBoost.

- SVM: El algoritmo Support Vector Machine es un clasificador generalmente usado para entrenar modelos supervisados en machine learning. Este puede ser implementado para modelos de clasificación y regresión. En cuanto a su funcionamiento, generalmente se clasifican los datos de entrenamiento en dos regiones creando un perímetro, conocido como hiperplano. Este hiperplano es único para cada modelo y es construido de una manera que reduce el error máximo y aumenta el margen geométrico. La clasificación de los datos de prueba es predicha calculando la distancia entre las características de los datos de entrenamiento y los de prueba. La clasificación que se le da a los datos de prueba es la misma que la del dato de entrenamiento más cercano (Varthakavi, Babu, & Reddy, 2019).

- Deep Learning: Se define como el conjunto de modelos y algoritmos computacionales que imitan la arquitectura de las redes neuronales biológicas. Cuando el cerebro recibe nueva información, este trata de compararlo con información previamente conocida para tratar de darle un sentido. El cerebro descifra la información y trata de asignarla a una o varias categorías. Este es el mismo concepto que emplea el Deep learning (Jakhar & Kaur, 2019).
- AdaBoost: Es la abreviación de Adaptive Boosting. Esta técnica es utilizada con otras técnicas de clasificación para mejorar el desempeño de los resultados. Primero, pesos uniformes son asignados a todos los ejemplos. Luego los pesos son ajustados de acuerdo con su clasificación, Los ejemplos incorrectamente clasificados se les asignan mayor peso en comparación a los correctamente clasificados (Rauf, Shahid, Ziauddin, & Safi, 2017).
- Árbol de decisiones: Este es un enfoque descendente usado para la clasificación, Este se basa en crear un cierto orden de atributos, de acuerdo con cuanto cada uno de estos ayuda a separar los datos en clases. Estos atributos están basados en información teórica y entre más información nos dé, más útil nos será. Entre las técnicas conocidas de este tipo están Random Forest, y AdaBoost Decision Stump.

Ahora que ya se han descrito las técnicas de extracción de características y clase, se comparan diversos métodos de detección de peatones con el uso de la *Tabla 2-1*, la cual es adaptada de una investigación con respecto al tema (Rauf, Shahid, Ziauddin, & Safi, 2017).

Tabla 2-1: Comparación de métodos de detección de peatones (Rauf, Shahid, Ziauddin, & Safi, 2017).

Método	Autor	Características	Clasificador	Porcentaje de falla (%)
1	Ouyang y Wang	HOG, flujo óptico	SVM	46.44
2	Chen	HOG, LBP	SVM	45.53
3	Zeng	HOG, CSS	Deep Learning	45.39
4	Yan	HOG	SVM	37.64
5	Benenson	HOG, LUV	SVM	34.81
6	Zhang	HOG, LUV	SVM	34.60
7	Benenson	HOG, flujo óptico	Árbol de decisiones	22.49
8	Zhang	HOG, LUV, flujo óptico	Ada Boost Decision Tree	17.1
9	Rauf	HOG, LUV, flujo óptico	Ada Boost Decision Stump	16.7

En la *Tabla 2-1* se puede observar que el método empleado por Rauf fue más efectivo que los métodos empleados previamente, aunque el método de Zhang tiene un porcentaje de falla muy similar al de Rauf. Los resultados de los métodos 1 a 6 se presentan en las *Figuras 2-2a* a *2-2f*. Los resultados del método 5, 6 y 8 no son mostrados en sus respectivas investigaciones.

De la tabla previamente mostrada se puede llegar a las siguientes observaciones:

- Se puede inferir que HOG es una técnica de extracción de características bastante usada en la detección de peatones, debido a que todos los investigadores en la tabla utilizaron este método, siendo la diferencia los métodos complementarios.
- Podemos observar que en cuanto a los clasificadores, los más exitosos han sido los que utilizan métodos basados en árbol de decisiones, esto se puede apreciar en la diferencia en cuanto al porcentaje de falla entre investigaciones que usan esta técnica y las que no, siendo esta diferencia de por lo menos 10%.

- Se puede ver que mientras las diversas investigaciones utilizan una variedad de extractores de características en sus estudios, la combinación de métodos de clasificación es menos común.



Figura 2-2a, 2-2b, 2-2c, 2-2d, 2-2e y 2-2f : Resultado del método 1, 2, 3, 4, 7 y 9 respectivamente (De izquierda a derecha) (Ouyang & Wang, 2013), (Chen, Ding, Xiao, & Han, 2013), (Zeng, Ouyang, & Xiaogang, 2013) (Yan, Zhan, Lei, & Lin, 2013), (Benenson, Omran, Hosang, & Schiele, 2014), (Rauf, Shahid, Ziauddin, & Safi, 2017).

2.1.2 Detectores de una etapa

Con el fin de procesar los resultados en menor tiempo y obtener una precisión suficiente, se proponen los detectores de una etapa. Este tipo de detectores ha logrado una mayor velocidad debido a que han mejorado la arquitectura de los modelos de dos etapas. Los algoritmos para tratar son SSD y las variantes de YOLO (Adarsh, Rathi, & Kumar, 2020).

YOLO v1:

YOLO v1 usa el *framework* de Darknet para redes neuronales y los datos de ImageNet-100 para entrenar este modelo. Su funcionamiento consiste en tomar una imagen y dividirla en una cuadrícula de $S \times S$ celdas. Cada celda es responsable de detectar 'n' objetos y encerrarlos en recuadros. Cada recuadro tiene 5 datos, los cuales son: el largo, el ancho, las coordenadas de su centro (en x, y) y el porcentaje de confianza. Luego, se calcula la probabilidad que cada celda contenga cada una de las clases en las que este modelo se encuentra entrenado. Finalmente, se le brinda los resultados de cada celda a la red neuronal convolucional, la cual obtiene las detecciones finales y les asigna la clase que le corresponde. Un ejemplo de la ejecución de este modelo se puede ver en la *Figura 2-3*. Esta versión cuenta con varias limitaciones. Una de estas es su incapacidad de detectar pequeños objetos dentro de grupos con muchos elementos. Otra desventaja es la dificultad que tiene en detectar objetos que tienen dimensiones diferentes a las presentes en los datos de entrenamiento (Adarsh, Rathi, & Kumar, 2020).

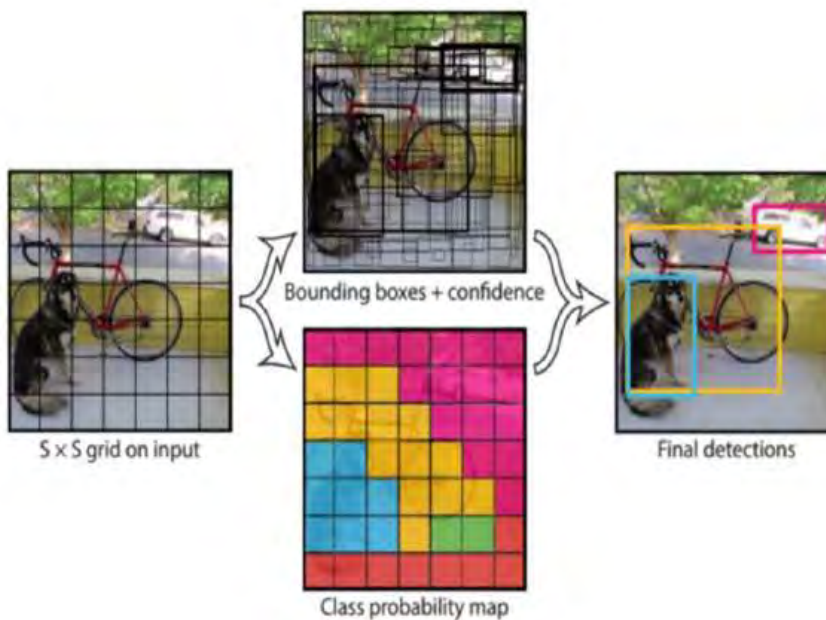


Figura 2-3: Ejecución del modelo YOLO. (Adarsh, Rathi, & Kumar, 2020).

YOLO v2:

Yolo v2 reemplaza a Yolo v1 ya que ofrece un adecuado balance entre tiempo de ejecución y precisión. Con el fin de mejorar la precisión, se utiliza normalización de lotes en cada capa de convolución. Además, el uso de características más finas ayuda a identificar objetos pequeños, modificando las capas de la red convolucional. En cuanto a velocidad, se utiliza el framework Darknet 19. Cabe resaltar que esta versión es capaz de reconocer 9000 categorías diferentes de objetos, por lo que este modelo también se conoce como YOLO 9000 (Adarsh, Rathi, & Kumar, 2020).

YOLO v3:

Esta variante de YOLO utiliza regresión logística para calcular el porcentaje de confianza de los objetos presente en cada recuadro. Además, es capaz de clasificar a un objeto con múltiples etiquetas ya que usa un clasificador logístico para cada una de las clases. YOLO v3 usa Darknet 53, el cual tiene 53 capas de convolución. Estas capas permiten un análisis más a profundidad comparado con el que permite YOLO v2 utilizando Darknet 19. Darknet 53 cuenta con filtros que son capaces de eliminar recuadros con porcentajes de confiabilidad menores a un umbral. Un segundo tipo de filtro es utilizado cuando múltiples recuadros están superpuestos entre sí, seleccionando el que mejor identifique al objeto, esto es conocido como NMS (Non maximum suppression). Otra ventaja que tiene sobre YOLO v2 es que se siguió mejorando el rendimiento ante objetos pequeños (Adarsh, Rathi, & Kumar, 2020).

SSD:

SSD o Single Shot Detector ejecuta el proceso de detección aplicando una sola vez un modelo basado en una red neuronal convolucional a una imagen. El procesamiento ocurre en las distintas capas de la red neuronal, donde cada capa funciona en un rango y escala

variable, utilizando distintos mapas de características, por esto se pueden detectar objetos de diferentes tamaños. Otra característica de este detector es que, a diferencia del YOLO, SSD no divide la imagen en una cuadrícula, sino que utiliza recuadros predefinidos, calcula el desfase de su posición y luego detecta objetos dentro de esos recuadros (Adarsh, Rathi, & Kumar, 2020).

2.2 Seguimiento de peatones

Una vez detectados los peatones, estos deben ser seguidos en el video, aplicando algoritmos de seguimiento de peatones. Muchas técnicas han sido propuestas ya sea como extractores de características o como técnicas de clasificación. En esta sección se describen los métodos más utilizados.

Filtro de partículas:

Esta solución se basa en estimar una región de búsqueda utilizando una gran cantidad de muestras, llamadas partículas, de estas se obtiene la frecuencia de un evento para obtener la probabilidad de este en el proceso de filtrado. La región de búsqueda es obtenida en base a las partículas con menor diferencia. La desventaja de este filtro es la pobre consistencia para seguir a múltiples peatones al mismo tiempo, ya que las partículas rápidamente migrarán hacia una persona y descartarán las demás (Quiong, 2015). Es por esto por lo que existen sub-algoritmos para obtener una región de búsqueda adaptativa. El ingeniero Liu Quiong propuso un método que determina una región de búsqueda aproximada, basándose en la velocidad del peatón, la cual en el contexto tráfico se asume uniformemente variable. El radio de la región de búsqueda puede ser obtenida prediciendo la velocidad del peatón en el fotograma actual teniendo como información la velocidad pasada y la aceleración. Otra manera en la cual Liu Quiong optimizó este método fue

mediante la introducción de cuadros delimitadores adaptativos, los cuales se escalan dependiendo de la distancia a la cámara, la cual se puede obtener conociendo la velocidad. Si se utilizara un cuadro delimitador estático, esta resultaría o muy grande, lo cual involucraría procesar mayor cantidad de fondo, o sería muy pequeña y no capturaría al peatón. Un ejemplo de su funcionamiento se puede en la *Figura 2-4*.



Figura 2-4: Funcionamiento del seguimiento usando filtro de partículas (*Gaddigoudar, Balihalli, Ijantkar, Iyer, & Maralappanavar, 2017*).

Modelo de Mezclas Gaussianas (GMM): Este método no se usa normalmente solo, sino como una ayuda para otros métodos de seguimiento debido a que ayuda a mejorar su estabilidad ya que apoya al establecimiento de un modelo para el fondo de la imagen, utilizando una mezcla de distribuciones gaussianas.

En este algoritmo los píxeles son tratados como un proceso aleatorio y expresados como un número k de distribuciones gaussianas. Debido a que el fondo cambia menos frecuentemente que el plano principal; el modelo del fondo es representado con distribuciones de mayor peso y menores varianzas. Las distribuciones son ordenadas de manera descendiente con respecto a la razón de su peso entre su varianza. En este método

se usa un umbral el cual nos ayuda a discernir el número de modelos gaussianos por los cuales está compuesto el fondo (Dong, Yan, & Wei, 2015).

Algoritmo del filtro de Kalman: Este método nos permite predecir el lugar más probable donde se podría encontrar un objeto en el presente, teniendo en cuenta el fotograma anterior. Luego, se busca la posición en el área vecina. Si el objeto se encuentra en esta área, el algoritmo continúa el proceso para el siguiente fotograma. La base del filtro es la predicción y actualización. (Dong, Yan, & Wei, 2015). Una desventaja de este filtro es su sensibilidad a la oclusión en comparación a otros métodos como el de partículas como se puede ver en las *Figuras 2-5a y 2-5b* (Gaddigoudar, Balihalli, Ijantkar, Iyer, & Maralappanavar, 2017).



Figura 2-5a: Implementación del filtro Kalman (Izquierda) y del filtro de partículas (Derecha) en casos sin oclusión. (Gaddigoudar, Balihalli, Ijantkar, Iyer, & Maralappanavar, 2017)



Figura 2.5b: Implementación del filtro Kalman (Izquierda) y del filtro de partículas (Derecha) en casos con oclusión. (Gaddigoudar, Balihalli, Ijantkar, Iyer, & Maralappanavar, 2017).

Algoritmo Camshift: Este algoritmo usa un modelo de histogramas de colores para poder seguir objetos. También es capaz de ajustar el tamaño de la ventana seguida y actualizar el modelo del objeto, mejorando la precisión de seguimiento. El resultado del fotograma anterior es tomado como valor inicial de la ventana de seguimiento del siguiente fotograma.

El proceso de este algoritmo consiste en establecer el modelo del objeto obteniendo el histograma de colores y luego obtener información de su posición, tamaño y orientación iterando de acuerdo con el modelo del objeto. Este algoritmo puede ser ayudado por el filtro Kalman, usando su predicción como coordenada de la primera iteración del algoritmo Camshift (Dong, Yan, & Wei, 2015) y sus resultados se pueden observar en la *Figura 2-6*.



Figura 2-6: Implementación de filtro Kalman y algoritmo Camshift (Dong, Yan, & Wei, 2015).

2.3 Conteo de peatones

La estimación de densidad de peatones predice el número de personas en un espacio determinado por una imagen. Los resultados de esta estimación pueden mejorar el desempeño en áreas relacionadas como detección y seguimiento (Jeong, Kim, & Shin, 2018).

Generalmente existen dos principales acercamientos para estimar el número de peatones. El primer tipo es el directo, y está basado en el seguimiento de personas individualmente, esto logra su propósito identificando cada transeúnte individual en las imágenes que se tienen como dato. El otro método, el indirecto, el cual extrae características locales u holísticas de los grupos de personas en las imágenes y aplica análisis de regresión o algoritmos de aprendizaje para estimar el número de peatones, en vez de tratar de contar cada transeúnte individualmente. Este tipo de procedimiento resulta ser más flexible, ya que no es necesario seguir cada peatón en la imagen (Huang, Chen, Yu, Yang, & Dong, 2018). Además, estos métodos tienden a ser computacionalmente más rápidos y robustos ante oclusiones parciales que pueden ocurrir en escenas llenas de gente (Jeong, Kim, & Shin, 2018). Entre los métodos usados para el conteo indirecto tenemos:

Método de los puntos de interés:

Estos extraen los puntos clave de una imagen para así estimar en número de personas utilizando la información de estos puntos. Se pueden dividir en dos enfoques diferentes: el basado en blobs, es decir, grupos de píxeles blancos unidos en una imagen en blanco y negro, y el basado en bordes. Los basados en blobs detectan los extremos locales de las respuestas a filtros como puntos claves. Entre los detectores de bordes se encuentran los métodos de Harris, FAST y BRISK. En una investigación publicada en la Conferencia Internacional sobre Convergencia de Tecnologías de Información y Comunicación de 2016 en que se comparan los diversos detectores de puntos clave (Jeong & Choi, A Comparison of Keypoint Detectors in the Context, 2016). Los tipos que fueron parte de la comparativa son explicados a continuación:

- SIFT: Es ampliamente reconocido como la alternativa estándar para detección de puntos clave invariantes en escala y rotación. Estos son hallados calculando la máxima/mínima diferencia de gaussianos (DoG).
- SURF: Esta es una de las modificaciones más populares del método anterior, esta disminuye su desempeño, pero aumenta su velocidad de procesamiento.
- Harris Corner: Este algoritmo es un método clásico para identificar esquinas utilizando información en escala de grises.
- FAST: Las esquinas son detectadas comparando la intensidad entre el píxel del centro y los que se encuentren en un anillo circular alrededor del centro. Este es un método frecuentemente usado para encontrar puntos clave en tiempo real debido a su eficiencia.
- BRISK: Es una variación multi escala del algoritmo FAST para encontrar puntos de interés. El detector BRISK exhibe una repetitividad equivalente al método SURF y es un orden de magnitud más rápido que este.

La comparación hecha en el estudio previamente mencionado tiene las siguientes consideraciones. Se utilizan los datos de PETS 2009. La implementación fue hecha utilizando la librería OpenCV, la cual contiene todos los métodos anteriormente explicados. El tiempo de procesamiento es calculado con una computadora personal con procesador Intel Xeon 3.0 GHz y Windows 7 como sistema operativo. (Jeong & Choi, A Comparison of Keypoint Detectors in the Context, 2016) Los resultados son mostrados la *Tabla 2-2*.

Tabla 2-2: Comparación entre detectores de puntos de interés (Jeong & Choi, *A Comparison of Keypoint Detectors in the Context*, 2016) (Adaptado).

Detector	Tiempo promedio de extracción de características (s)	Tiempo promedio de emparejamiento (s)	Tiempo total (s)	Exactitud (%)
Harris	10.38	5.45	15.83	78.62
FAST	2.31	17.41	19.72	82.95
BRISK	11.28	4.32	15.6	89.95
SIFT	213.15	7.02	220.17	84.71
SURF	41.93	9.86	51.8	71.27

De esta tabla se puede concluir que estos resultados experimentales muestran que el detector BRISK es el más adecuado para el conteo de peatones, debido a que cuenta con el mayor porcentaje de exactitud y el menor tiempo de procesamiento total, aunque cabe aclarar que el método más rápido para extraer características es el FAST, ya que puede hacer esta tarea 5 veces más rápido que el BRISK. Otra observación importante es que en la mayoría de los casos la tarea de extracción de características toma mayor tiempo que la de emparejamiento, siendo esta diferencia en el caso de BRISK de 3 a 1. Una ejemplificación de los puntos de interés obtenidos por BRISK y SURF se puede ver en la *Figura 2-7*.



Figura 2-7: Ejemplo de detección de los puntos de interés (verdes SURF y rojos BRISK) (Jeong, Kim,, 2018).

Análisis Temporal y Espacial:

Este método crea una matriz espacio temporal que contiene la posición de los peatones y el tiempo en que aparecieron. Una sustracción del fondo es aplicada para detectar a los objetos móviles para poder crear esta matriz. En la *Figura 12* se muestra la línea que se traza para obtener la matriz espaciotemporal, la cual debe ser perpendicular al flujo. Esta línea funciona como límite que define donde se va a recolectar la información del plano principal. La línea auxiliar (con guiones) mostrada en la *Figura 2-8* es utilizada para ayudar al sistema a encontrar la dirección de los peatones (Yu, Gong, Yang, & Li, 2014).



Figura 2-8: Línea principal y línea auxiliar (Yu, Gong, Yang, & Li, 2014).

En cada fotograma, se marca el plano principal en blanco y el fondo en negro. La matriz espaciotemporal es un conjunto de líneas ordenadas cronológicamente las cuales son iguales a las líneas horizontales más cercanas al límite pertenecientes a los fotogramas binarizados. La matriz obtenida por la línea principal y auxiliar durante el fotograma 1 a 200 se muestra en la *Figura 2-9* En este caso las matrices son obtenidas por cada ciclo de video (Yu, Gong, Yang, & Li, 2014).

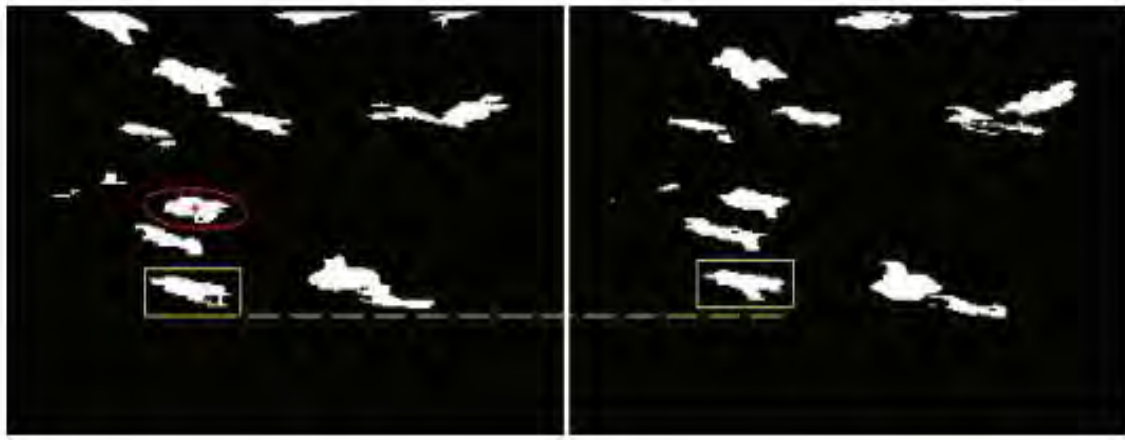


Figura 2-9: Matrices espaciotemporales: izquierda: línea principal, derecha: línea auxiliar (Yu, Gong, Yang, & Li, 2014)

Luego de obtener la matriz espacio temporal, se utilizan clasificadores para distinguir cada región en blanco en dos clases, un peatón y más de un peatón. En caso de ser más de un peatón se utiliza el método de *mean-shift clustering* para calcular la cantidad de personas, este es un método de agrupamiento muy similar al algoritmo Camshift explicado previamente, siendo la diferencia que este solo aplica para una imagen (Dong, Yan, & Wei, 2015). Los resultados del método se ven en la *Figura 2-10*, en este los puntos verdes son los peatones individuales y los puntos de otros colores son los centros de los grupos de personas. La información sobre su efectividad y tiempo de procesamiento se pueden ver en la *Tabla 2-3* (Yu, Gong, Yang, & Li, 2014). Se tienen las siguientes consideraciones, la primera es que se utilizó una base de datos creada por los investigadores, así como la base de datos de UCSD (University of California San Diego). La segunda consideración es que los resultados conseguidos se obtuvieron utilizando el programa Matlab en una computadora con procesador Pentium Dual Core de 3.2 GHz y memoria 2G.

Teniendo en cuenta los resultados de la *Tabla 2-3* y los de la *Tabla 2-2* se puede observar que el método de análisis espacio temporal tiene mejor tiempo de procesamiento y efectividad que los métodos de puntos de interés, pero cabe resaltar que debido a que no utilizan la misma base de datos ni el mismo equipo de procesamiento, estos resultados no son totalmente concluyentes.

Tabla 2-3: Resultados del proceso (Yu, Gong, Yang, & Li, 2014).

Tiempo de procesamiento	Efectividad (%)
4.04(s)	95

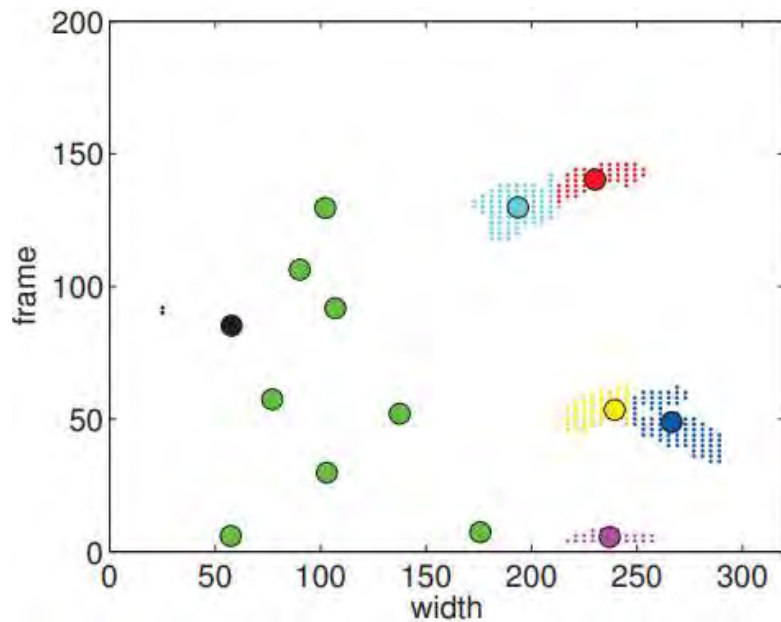


Figura 2-10: Resultados del análisis (Yu, Gong, Yang, & Li, 2014).

2.4 Plataformas móviles para cámaras

En la presente sección se presentarán diferentes diseños de plataformas móviles presentes en investigaciones, las cuales son utilizadas para mantener la cámara direccionada hacia el objetivo deseado, lo cual en la mayoría de los casos involucra mantenerlo en el centro del campo de visión de la cámara. Estas se presentan a continuación.

Sistema Pan-Tilt y Pan-Tilt-Zoom (PTZ)

Se define a un sistema Pan-Tilt como un conjunto, que cuenta con una cámara, una unidad estacionaria, un miembro intermedio entre ambas partes y dos articulaciones rotatorias. La primera articulación conecta la cámara al miembro intermedio y la segunda articulación conecta el miembro intermedio con la unidad estacionaria (Estados Unidos Patente n° US9007474B2, 2009). Estas articulaciones le dan la capacidad al sistema de rotar en la dirección horizontal y vertical rápidamente. (Ukida, 2010) En el caso del sistema PTZ, este cuenta con los mismos elementos del Pan-Tilt, pero se le añade la

capacidad de controlar el zoom de la cámara. En la *Figura 2-11* se muestran ejemplos de este tipo de sistema.

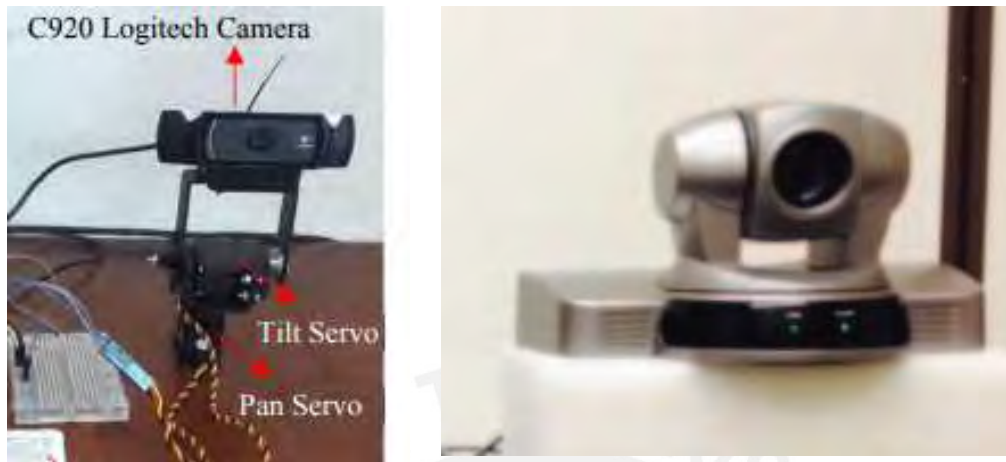


Figura 2-11: Ejemplos de sistemas Pan Tilt y PTZ de izquierda a derecha (Rauf, Shahid, Ziauddin, & Safi, 2017), (Lee & Zixiang, 2015)

Uso de impresión 3D

En la investigación de Cynthia Sung, se propone la fabricación de mecanismos y robots *Print-and-fold* que cuenten con articulaciones mediante impresión 3D. Estos consisten en imprimir el patrón de plegado en una plancha y doblarla para llegar a su forma final, como se ven en la *Figura 2-12a*. Los sensores y actuadores se pueden incorporar al robot durante el proceso de fabricación, imprimiendo circuitería y colocando los componentes en el patrón de plegado antes del doblado. Todas las articulaciones proveen naturalmente un espacio para los actuadores y circuitos. Para demostrar el funcionamiento de este método de fabricación, se desarrollaron diferentes mecanismos, entre ellos una plataforma móvil para cámara de smartphone. Esta se muestra en la *Figura 2-12b*.

El sistema presentado previamente, contaba con dos grados de libertad, teniendo posibilidad de rotar en el eje vertical y horizontal gracias a que cuenta con una articulación

de bisagra y otra de pivote. La actuación del sistema es realizada independientemente por dos servos los cuales eran controlados externamente.

La articulación de pivote estaba diseñada para poder rotar 1.5π radianes en el eje vertical, mientras que la bisagra estaba diseñada para rotar $2/3 \pi$ radianes el eje horizontal, pero debido a la carga del celular la bisagra tuvo que ser limitada a un giro de 0.5π radianes; a causa que más de la mitad del peso del dispositivo estaba sobre la bisagra (la plataforma pesaba 91 g y el smartphone 116 g), la plataforma tendía a doblarse cuando se intentaba rotar ángulos mayores a los establecidos.

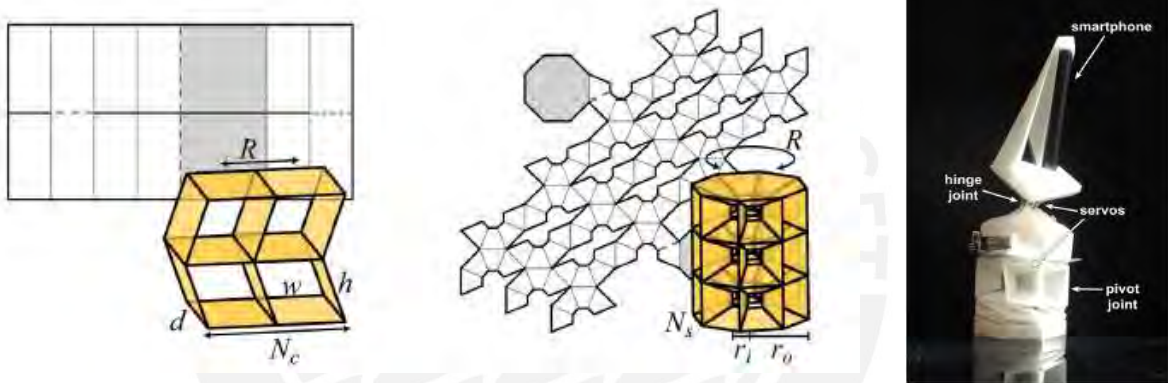


Figura 2-12a y Figura 2-12b: Ejemplo de articulaciones Print and fold y Plataforma de cámara impresa en 3D (Sung & Rus, 2015).

Brazo Robótico y cámaras PTZ:

En las investigaciones del ingeniero Ukida, se propuso un sistema de seguimiento de objetos que utiliza cámaras Pan-Tilt y un brazo robótico. Esto se debe a que las cámaras Pan-Tilt tienen rangos de rotación limitados mientras que los brazos robóticos tienen mayor rango de rotación, pero velocidades menores. Por esto se planteó este sistema que tendría una alta velocidad y un amplio rango para el seguimiento de objetos.

En la *Figura 2-13* se puede observar el sistema propuesto, el cual consta de un brazo robótico y dos cámaras pan-tilt las cuales son ubicadas en la parte superior del brazo. El robot cuenta con 6 grados de libertad, pero solo se utilizan dos articulaciones, para lograr

rotación en el eje horizontal y vertical. La velocidad angular de cada articulación es de 25 grados por segundo y cuenta con una rotación máxima de 30 grados en el eje horizontal y 60 grados en el vertical. En el caso de las cámaras, estas están separadas por una distancia de 200 mm. Su velocidad de rotación es de 45 grados por segundo y su rotación máxima es de 320 grados en el eje vertical y de 90 grados en el eje horizontal (Ukida, 2010).

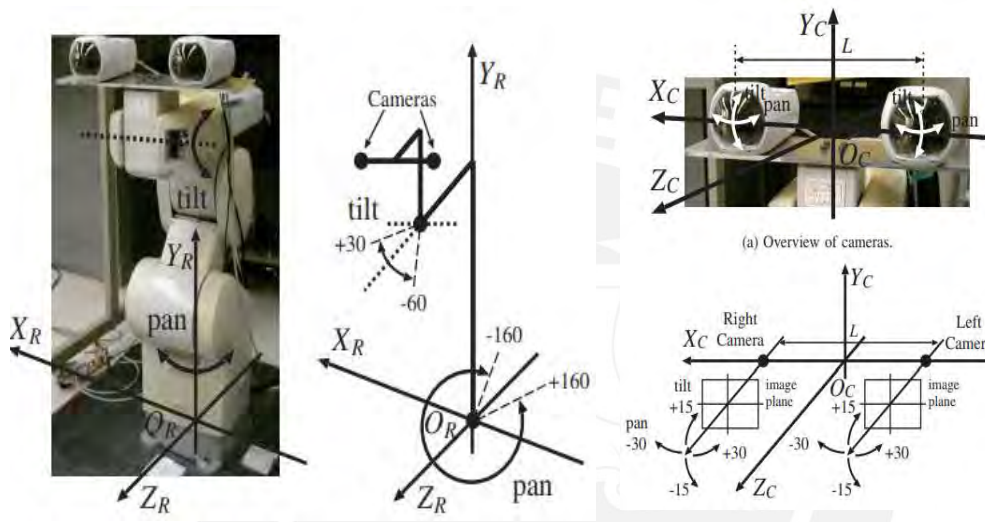


Figura 2-12 Brazo robótico, cámaras PTZ con sus respectivos sistemas de coordenadas (Ukida, 2010).

2.5 Control de plataformas móviles

En esta sección se discutirán los métodos emergentes para el control de plataformas móviles.

Control por seguimiento:

Este método se apoya del algoritmo de seguimiento para poder cambiar la posición de la cámara de tal manera que el objeto de interés se mantenga en el centro de la cámara y ajusta el zoom para mantener su tamaño constante. (Lee & Zixiang, 2015). Los resultados del presente método se observan en la *Figura 2-14*.



Figura 2-13 Demostración del control por seguimiento . (Lee & Zixiang, 2015).

Deep Q-Network (DQN):

Este es un método que permite aprender normas de comportamiento exitosas a través de entradas sensoriales utilizando aprendizaje de reforzamiento (RL). El aprendizaje de reforzamiento es un método típico de aprendizaje en machine learning; este es un proceso de prueba y error en el cual un agente explora y explota un espacio de posibles estrategias, y a través de realimentación en forma de recompensas y sanciones, encuentra el patrón con mayor recompensa. En el caso del DQN, se trata de optimizar el valor de Q, el cual está basado en una función basada en la ecuación de Bellman, siendo este valor relacionado con la recompensa que se busca. En este caso, los valores generados son utilizados para controlar la plataforma móvil y entre mayor sea la aproximación a la posición requerida, mayor será el estímulo positivo recibido. (Kim, Kim, & Sungjoo, 2019). Los resultados de este método se pueden apreciar en la *Figura 2-15*.

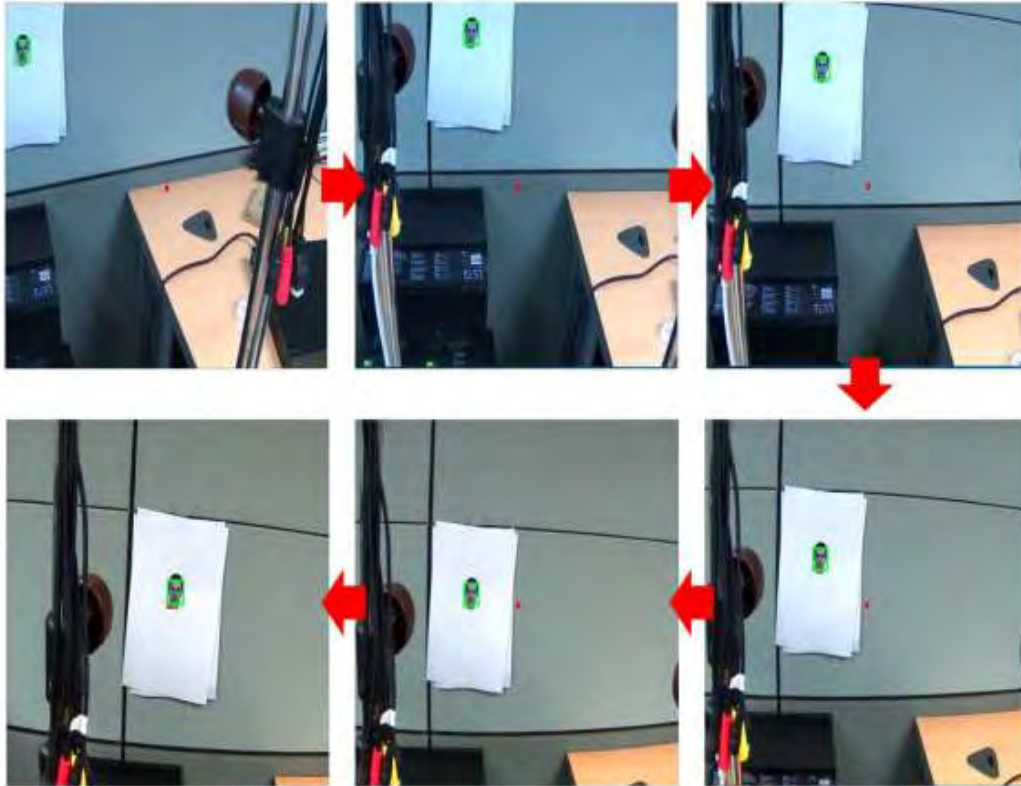


Figura 2-14: Resultados del control de acuerdo con un objeto móvil (Kim, Kim, & Sungjoo, 2019).

Utilización de teoría de control clásico

Se han realizado estudios acerca de la utilización de compensadores de atraso y adelanto, así como algoritmos PID para el control de cámaras móviles destinadas al seguimiento facial. (Yosafat, Machbub, & Hidayat, 2017). El sistema de control utilizado por Yosafat en su investigación tiene como entrada(referencia) los centroides de la cara a seguir en cada fotograma del video. Las coordenadas de los datos de salida son utilizadas para controlar el sistema, en este caso un sistema Pan-Tilt. Estas le permiten moverse a la posición deseada automáticamente. La salida es realimentada a la entrada para así obtener el error. Dependiendo del error, el algoritmo generará una señal PWM para mantener el objetivo en el centro de la imagen. El diagrama de bloques de este sistema se puede apreciar en la *Figura 2-16*.

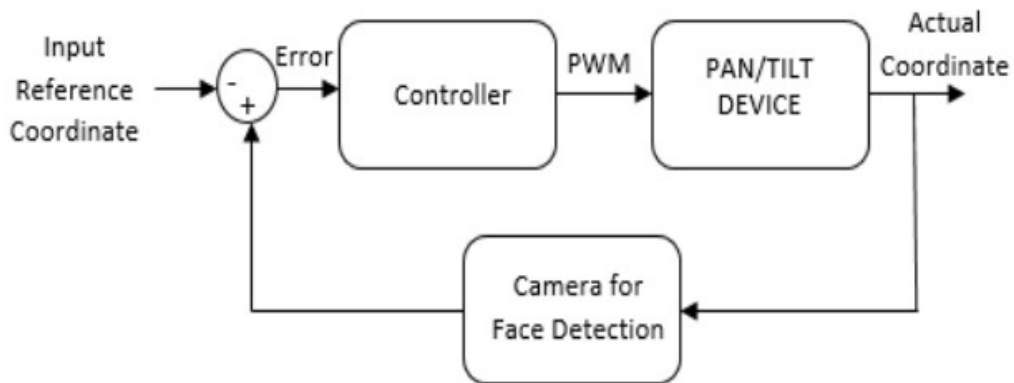


Figura 2-156 Diagrama de bloques del sistema de reconocimiento facial implementado (Yosafat, Machbub, & Hidayat, 2017)

El modelamiento de esta planta fue hecho de manera experimental utilizando MATLAB. En este caso, la función que describe la dinámica del sistema se ajusta al modelo de validación un 88.6% de las veces en con respecto eje horizontal y un 85.65% con respecto al vertical. Esta estimación fue utilizada para el diseño de los controladores teniendo en consideración que el tiempo de pico debe ser menor a 0.5 segundo, y el sobre impulso no debe exceder el 10% (Yosafat, Machbub, & Hidayat, 2017).

Para el diseño del compensador de atraso-adelanto, se utilizó el método SISOTOOL para obtener el modelo recursivo del compensador y así poder implementarlo en un algoritmo el cual tome los valores previos del error para ejecutar la acción de control. En el caso del controlador PID, se utilizó el método de Ziegler Nichols, el cual determina los parámetros de control de manera empírica para posteriormente poder ser implementados.

Los resultados de los experimentos se pueden ver a continuación en las Figuras 2-17, 2-18.



Figura 2-167: Respuesta de la plataforma a una cara en movimiento (Yosafat, Machbub, & Hidayat, 2017).

Como se puede observar en la Figura 2-17, tanto el compensador como el controlador PID tienen el mismo error en estado estable, siendo este de aproximadamente ± 2 píxeles en una imagen de resolución 640 x 480 píxeles lo cual es considerado un error aceptable y le permite al sistema ejecutar el seguimiento facial exitosamente como es visto en la Figura 2-17.

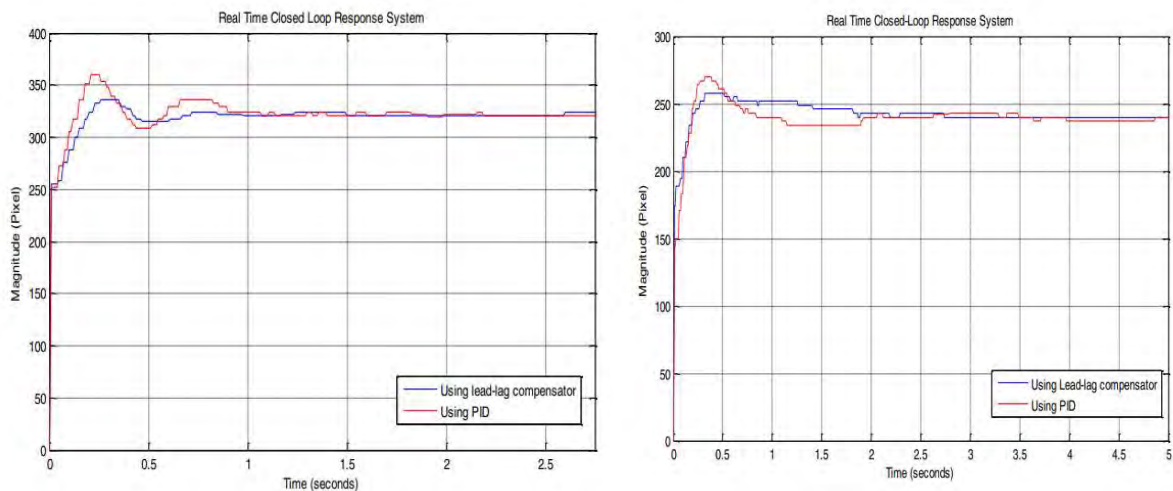
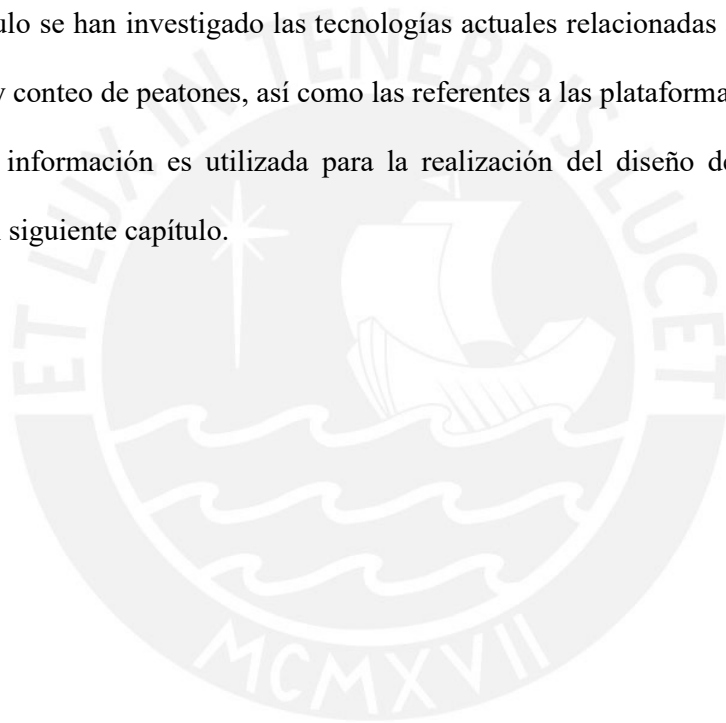


Figura 2-17: Respuesta en lazo cerrado del eje horizontal y vertical respectivamente (Yosafat, Machbub, & Hidayat, 2017).

De la Figura 2-18, la cual describe la respuesta en el tiempo del sistema en ambas coordenadas, se puede observar que el sistema que utiliza un controlador PID muestra mayor cantidad de oscilaciones en comparación al que usa un compensador de atraso-adelanto. Asimismo, se puede afirmar que su error en estado estable es el mismo ya que sus respuestas tienden a converger en el mismo punto. Otra observación es que el tiempo

de establecimiento, es decir el tiempo que se demora el sistema en llegar a la referencia, es menor en el sistema con compensador en comparación con el sistema que cuenta con un controlador PID. Teniendo estas observaciones en consideración, se puede concluir, que de acuerdo con los resultados brindados por la investigación de los ingenieros Yosafat, Machbub y Hidayat, el controlador más conveniente en esta situación es el compensador atraso-adelanto, debido a que cuenta con menor oscilación y tiempo de establecimiento con respecto al controlador PID.

En este capítulo se han investigado las tecnologías actuales relacionadas a la detección, seguimiento y conteo de peatones, así como las referentes a las plataformas móviles y su control. Esta información es utilizada para la realización del diseño de la propuesta presente en el siguiente capítulo.



Capítulo 3 Diseño de la propuesta

En el presente capítulo se examinan los aspectos relacionados con el diseño de la propuesta, entre estos se abordará la elaboración de la lista de requerimientos, la estructura de funciones, la matriz morfológica, el concepto de solución óptimo, entre otros.

Se está considerando un sistema que cuente con una cámara móvil, este es ubicado en el poste de una esquina de tal manera que analice cruces peatonales semaforizados adyacentes (a 90 grados). El sistema se direccionará al cruce peatonal respectivo automáticamente teniendo en cuenta el video grabado. Un esquema del posicionamiento del sistema se puede ver en la *Figura 3-1a* mientras que las potenciales perspectivas del sistema con respecto a los cruces se observan en las *Figuras 3-1b y 3-1c*.

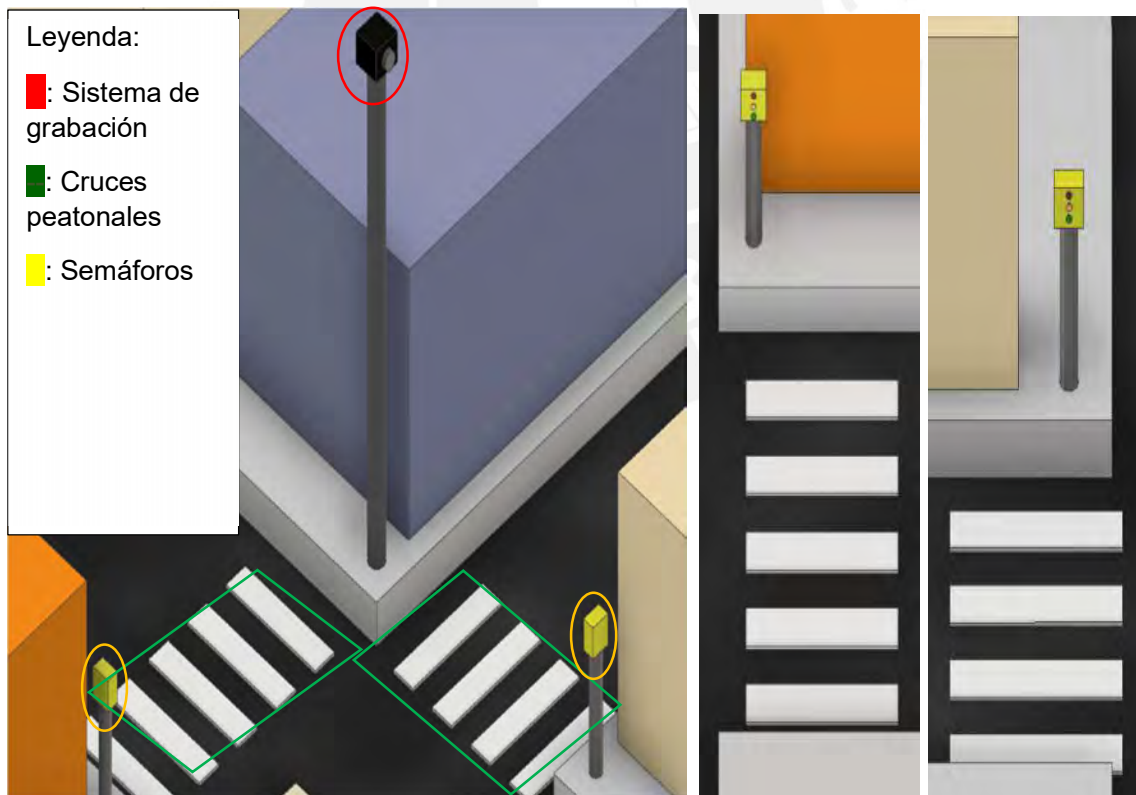


Figura 3-1,3.1b y 3.1c: Esquema del posicionamiento del sistema, perspectivas del sistema con respecto a los cruces peatonales (respectivamente de izquierda a derecha). Elaboración propia.

3.1 Lista de requerimientos

En la *Tabla 3-1* se presenta la lista de requerimientos del sistema.

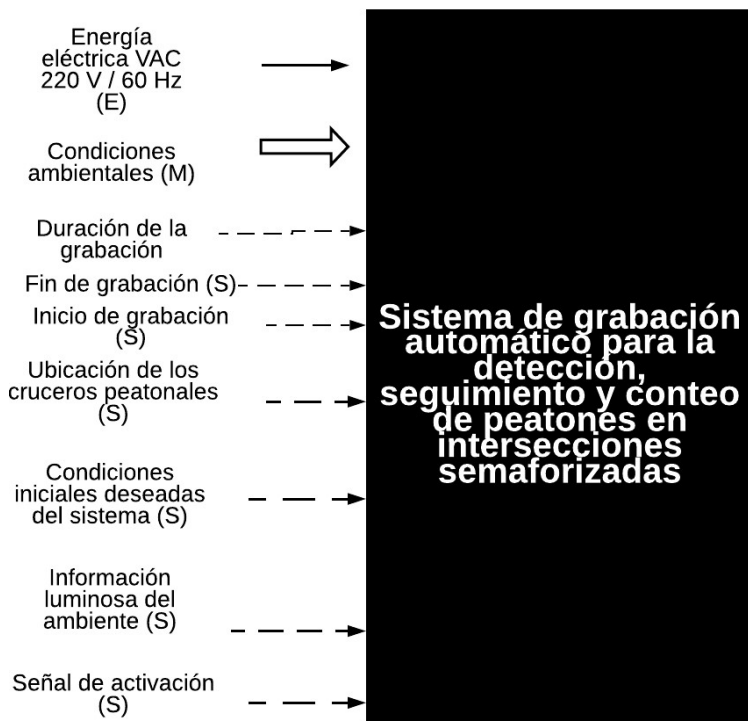
Tabla 3-1 Lista de requerimientos. Elaboración propia.

Lista de requerimientos	
Tema	Diseño de sistema de grabación automático para la detección, seguimiento y conteo de peatones en intersecciones semaforizadas
Deseo(D) o Exigencia(E)	Descripción
E	Función principal: El sistema graba automáticamente el flujo peatonal en intersecciones, obteniendo datos sobre el volumen de tráfico peatonal.
E	Geometría: El sistema cuenta con unas dimensiones máximas de 25x25x25 cm.
E	Control: Se controla la posición de la cámara de tal manera que se dirccione a los peatones automáticamente.
E	Señales: El sistema detectará peatones que se muevan cruzando las intersecciones viales con una resolución apropiada para su procesamiento.
E	Software: El sistema procesa el video grabado, detecta peatones, los sigue y cuenta.
E	Seguridad: El sistema cuenta con una protección IP 54.
E	Cinemática: El sistema es capaz de hacer seguimiento a los peatones
E	Montaje: El sistema permanece fijo sobre un poste. Contará con la capacidad de ser desmontado para mantenimiento.
E	Software: El conteo es efectivo en más del 90% de las veces.
D	Software: Se identificarán los cruces peatonales automáticamente.
E	Entrada: El sistema captura video durante el día y a la altura de un poste siendo la inclinación de la cámara de 90 a 65 grados (Yu, Gong, Yang, & Li, 2014).
E	Energía: El sistema utiliza baterías para su funcionamiento.
D	Software: Los resultados del conteo se comprueban utilizando una cámara estática situada de tal manera que se puedan observar los dos cruces peatonales al mismo tiempo.

3.2 Caja negra del sistema

En esta sección se muestra la caja negra del sistema de grabación automático para la detección, seguimiento y conteo de peatones en intersecciones semaforizadas. Esto nos ayuda a delimitar su función principal, así como sus entradas y salidas, sin considerar ningún detalle interno. En la *Figura 3-2* se muestra dicho diagrama, la función principal, las entradas y salidas, las cuales son explicadas a continuación:

ENTRADAS



SALIDAS

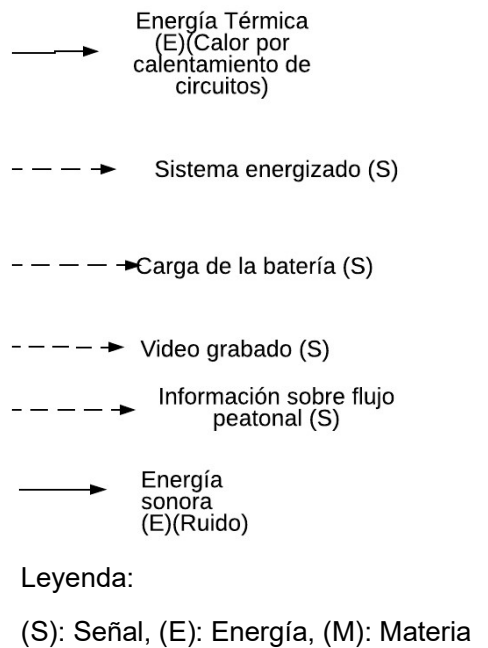


Figura 3-2: Caja negra del sistema. Elaboración propia.

3.2.1 Función principal

El sistema tiene como función grabar automáticamente en intersecciones semaforizadas, detectando, siguiendo y contando transeúntes, así obteniendo datos sobre el volumen de tráfico peatonal.

3.2.2 Entradas y salidas

El sistema cuenta con diversas entradas y salidas de energía, materia y señal, la cuales son detalladas a continuación.

3.2.2.1 Entradas de energía

Considerando la disponibilidad de la energía eléctrica a 220 VAC 60 Hz provista por la red eléctrica peruana, se decide que usar este tipo de energía para cargar las baterías del sistema.

3.2.2.2 Entradas de materia

Tomando en cuenta la premisa que el sistema estará sujetado en un poste y a la intemperie por prolongados periodos de tiempo, las condiciones ambientales se tienen presente como entradas del sistema.

3.2.2.3 Entradas de señal

Estas señales contienen la información a procesar, así como la configuración inicial del sistema y las ordenes necesarias para el funcionamiento y control del sistema. Entre estas se tiene la parada de emergencia, inicio de grabación, fin de grabación, duración de grabación, ubicación de cruces peatonales, condiciones iniciales deseadas del sistema, señal de activación e información luminosa del ambiente.

El sistema se energiza o desenergiza con la señal de activación, luego es necesario configurarlo, indicándole la posición de los cruces peatonales a monitorear, así como la posición inicial del sistema y la duración de la grabación. Luego, es necesario enviar una señal de inicio de grabación para iniciar el proceso, en el cual se adquiere la información luminosa del ambiente en forma de video para obtener datos acerca del flujo peatonal. Finalmente, este proceso se detiene después de su duración programada o mediante una señal de final del proceso.

3.2.2.4 Salidas de energía

En las salidas de energía se considera la energía térmica emitida por los circuitos del sistema y el ruido provocado por el movimiento de los actuadores. Estas son consideradas como pérdidas, las cuales son comunes en toda máquina.

3.2.2.5 Salidas de señales

En esta sección se encuentran señales que son utilizadas para conocer aspectos importantes del sistema, como saber si este se encuentra energizado e información acerca de la batería. Dichas salidas nos permiten conocer si la máquina está en funcionamiento y cuando necesitará un cambio de baterías.

Además, teniendo en cuenta los objetivos del sistema, este nos permitirá tener una base de datos de videos de intersecciones ofreciendo como salida el video grabado. Asimismo, para cumplir su función principal, otra salida a considerar es la información del flujo peatonal obtenida, en esta se considera el número de peatones por ciclo semafórico, así como otras características importantes como la velocidad promedio de los peatones.

3.3 Estructura de funciones

En esta sección se presenta la estructura de funciones del sistema, así como los subsistemas y funciones que este posee. La estructura de funciones se puede ver en la *Figura 3-3* y nos permite indicar como se comunica cada función de cada subsistema a través de líneas que facilitan la visualización del proceso. En este caso nos ayuda a explicar el proceso de grabación automática del sistema y como las entradas y salidas explicadas previamente interactúan con este.

Estructura de funciones

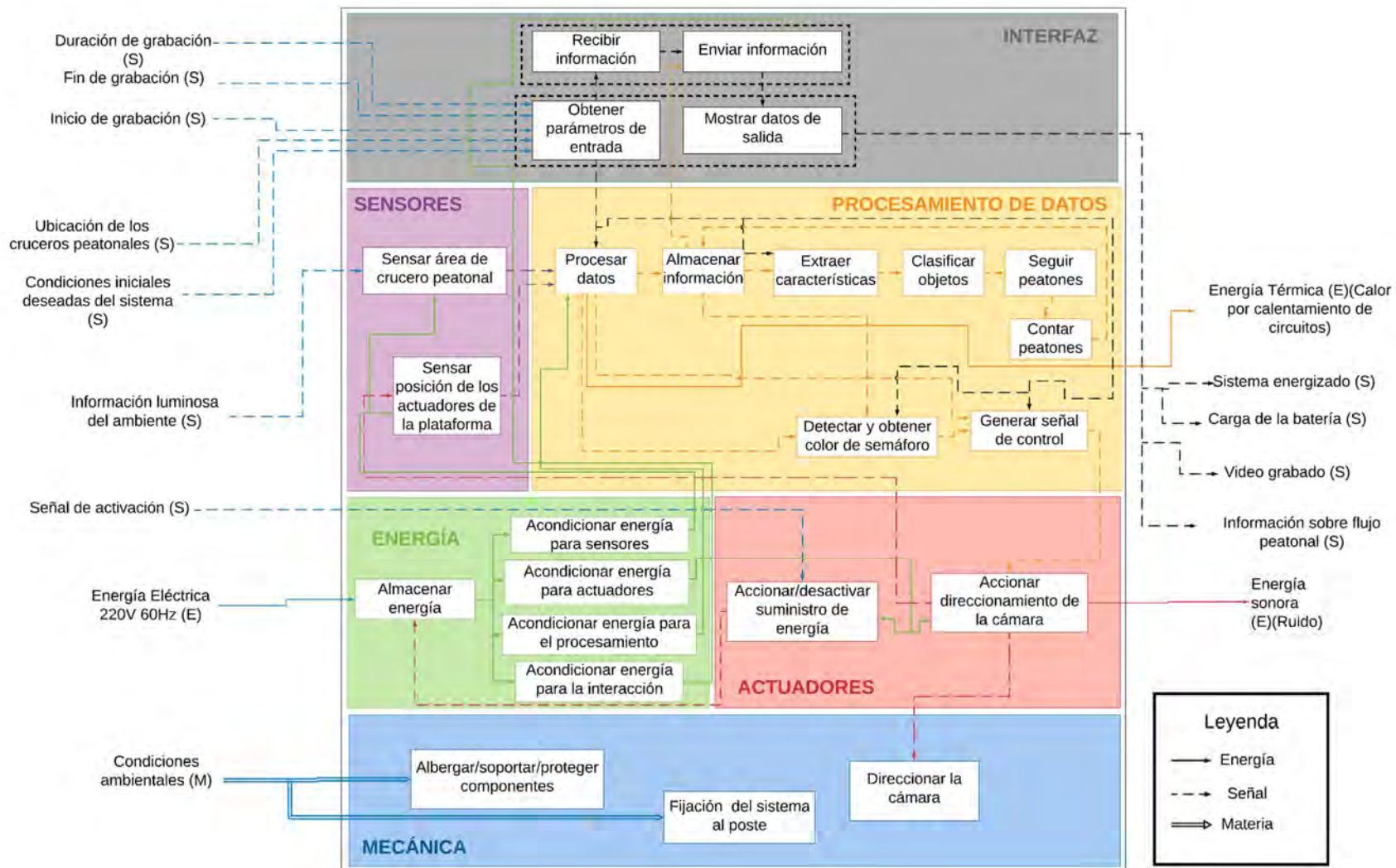


Figura 3-3: Estructura de funciones. Elaboración propia.

3.3.1 Subsistemas

Los subsistemas han sido seleccionados de tal manera que se pueda comprender del sistema y dividirlo efectivamente. Los subsistemas elegidos han sido: subsistema de interfaz, subsistema de sensores, subsistema de procesamiento de datos, subsistema de energía, subsistema de actuadores y subsistema mecánico.

3.3.1.1 Subsistema de interfaz

Este subsistema permite la interacción con el operador, permitiendo al sistema obtener las señales de entrada para su correcto funcionamiento como las señales de inicio y fin del proceso, además permite al usuario darle al sistema una configuración inicial. Este subsistema también es capaz de brindar las señales de salida al usuario las cuales son las condiciones energéticas del sistema y los datos de flujo peatonal. Debido a los requerimientos, los cuales especifican que el sistema deberá ser montado en un poste, la interfaz deberá ser remota ya que esto facilitaría la interacción con el usuario, es por esta razón que se incluye también la función de comunicación entre la interfaz y el sistema. Las funciones de este subsistema son explicadas a continuación.

- Obtener parámetros de entrada y mostrar datos de salida: Esta función permite al usuario realizar distintas tareas de manera remota, como indicar la configuración inicial del sistema, especificando los cruces peatonales y semáforos a grabar, la posición inicial de la cámara, entre otros; iniciar y finalizar el proceso. Además, el operario es capaz de saber si el sistema está energizado, conocer el estado de carga de la batería y visualizar los datos relacionados con el flujo peatonal, entre estos se considera la cantidad de peatones por ciclo semafórico, la velocidad promedio de los peatones, entre otros datos relevantes. En la *Figura 3-3*, se muestra a esta función dividida en dos bloques, ubicadas dentro recuadro

punteado, esto se hizo para hacer más entendible el flujo de información entre los bloques.

- Recibir y enviar información: La presente función permite conectar la interfaz remota con el resto del sistema, recibiendo los datos de entrada de este, así como enviando los datos de salida que son mostrados por la interfaz externa. Al igual que la función anterior, en la *Figura 3-3*, esta función fue dividida en dos para hacer más entendible el flujo de datos.

3.3.1.2 Subsistema de sensores

Este subsistema se encarga de obtener la información interna y externa necesaria para poder realizar el proceso. En el caso de la información interna, se sensorará la posición de los actuadores de la plataforma, y en el caso de la externa, se conseguirá la información luminosa del área donde se encuentran los cruces peatonales.

- Sensor área de cruces peatonales: El subsistema es capaz de captar la información luminosa del ambiente en forma de video de las regiones donde se encuentren los cruces peatonales siendo esta el área donde se pueden observar los peatones y semáforos.
- Sensor posición de los actuadores de la plataforma: El subsistema puede determinar la posición de los actuadores encargados del movimiento de la plataforma para así conocer si la cámara está siendo direccionada correctamente.

3.3.1.3 Subsistema de procesamiento de datos

Este subsistema es el encargado de procesar el video y posiciones obtenidas para poder controlar el movimiento de la plataforma y conseguir información del flujo peatonal en las intersecciones indicadas. Esto se puede lograr realizando las funciones explicadas a continuación.

- Procesar datos: Esta función se refiere al componente físico responsable del procesamiento del video capturado y el control de la plataforma móvil.
- Almacenar información: Esta función permite al sistema almacenar el video grabado para su posterior uso y para ser parte de una base de datos de intersecciones. Asimismo, se almacena información relevante como el número de peatones por ciclo semafórico, así como su velocidad promedio, entre otras características.
- Detectar y obtener color del semáforo: El subsistema es capaz de reconocer el semáforo en el video grabado dada la zona aproximada del mismo, así como reconocer su color para determinar el ciclo semafórico.
- Extraer características: Después de almacenado el video relacionado a la al cruce peatonal, este se procesa para detectar los peatones en la zona de interés, como primer paso se extraen las características de los objetos encontrados.
- Clasificar objetos: Después de obtener las características de los objetos, se procede a clasificarlos para determinar si son peatones.
- Seguir peatones: Se predice la trayectoria de los peatones teniendo en cuenta fotogramas anteriores para así poder distinguir cada peatón.
- Contar peatones: Tomando en cuenta el video grabado y procesamiento previo, se cuenta el número de peatones que pasan cada ciclo semafórico.
- Generar señal de control: Se toma en cuenta en cuenta la posición previa de la cámara, el color del semáforo y la ubicación de los cruceos peatonales para determinar el movimiento de los actuadores relacionados con el direccionamiento de la cámara.

3.3.1.4 Subsistema de energía

Este subsistema almacena y provee energía para mantener al sistema en funcionamiento continuo. Para el subsistema en mención se tienen las siguientes funciones.

- Almacenar energía: Se almacena la energía para el correcto funcionamiento del sistema.
- Acondicionar energía para los sensores: Se acondiciona la energía suministrada por las baterías a la necesitada por el subsistema de sensores.
- Acondicionar energía para los actuadores: Se acondiciona la energía suministrada por las baterías a la necesitada por el subsistema de actuadores.
- Acondicionar energía para el procesamiento: Se acondiciona la energía suministrada por las baterías a la necesitada por el subsistema de procesamiento.
- Acondicionar energía para la interacción: Se acondiciona la energía suministrada por las baterías a la necesitada por el subsistema de interacción.

3.3.1.5 Subsistema de actuadores

Este subsistema se encarga del accionamiento de los mecanismos que hacen posible el funcionamiento del sistema. Se considera el direccionamiento de la cámara y el suministro de energía.

- Accionar/desactivar el suministro de energía: Esta función se encarga de permitir o impedir a las baterías suministrar energía a la máquina dependiendo de la señal de energización o de la parada de emergencia.
- Accionar direccionamiento de la cámara: Se acciona el mecanismo encargado de mover la plataforma para así direccionar la cámara encargada de grabar peatones automáticamente.

3.3.1.6 Subsistema mecánico












Este subsistema cuenta con la estructura responsable de la dinámica del sistema y de la interacción de este con el ambiente, las funciones consideradas son el direccionamiento de la cámara, la protección de los componentes del sistema y la fijación del sistema al poste. Estas son explicadas a continuación.






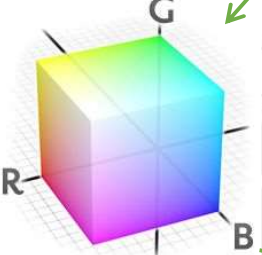
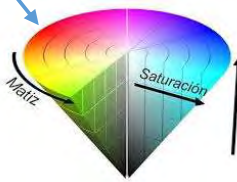

- Albergar/soportar/proteger componentes: Debido a que el sistema se encuentra a la intemperie por periodos prolongados de tiempo, existe la necesidad de albergar los componentes para poder protegerlos de las condiciones ambientales.
- Fijación del sistema al poste: El sistema debe sujetarse al poste donde se encuentra para poder realizar sus operaciones correctamente.
- Direccionar la cámara: El subsistema cuenta con una plataforma móvil la cual es utilizada para direccionar la cámara hacia el flujo peatonal.













3.4 Matriz Morfológica



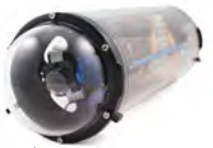


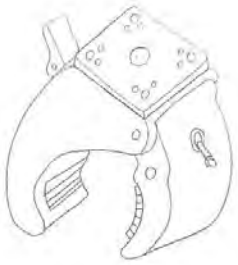
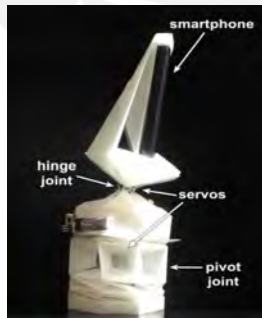

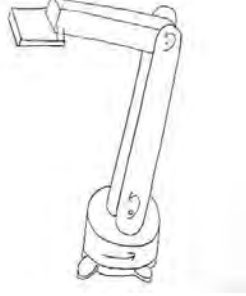
La matriz morfológica presente en la *Tabla 3-2* muestra principios de solución para cada función detallada en la sección 3.3.1. Cada principio, de ser posible, ha sido descrito por un dibujo o figura referencial al nombre del mismo principio. Asimismo, se señalan mediante flechas de diferentes colores tres conceptos de solución. El primer concepto de solución utiliza flechas de color azul, mientras que el segundo concepto utiliza flechas verdes y el tercero usa flechas de color amarillo.

Tabla 3-2 Matriz Morfológica. Elaboración propia.

Subsistema	Función	Opción 1	Opción 2	Opción 3
Interfaz	Obtener parámetros de entrada y mostrar datos de salida	Interfaz para computadora. 	App para celular 	Control remoto con pantalla 
	Recibir y enviar información	Antena Wifi 	Antena Bluetooth 	Antena Wimax 
Sensores	Sensar área de cruceo peatonal	Cámara embebida 	Cámara digital 	Cámara analógica 
	Sensar posición de los actuadores	Encoder 	Acelerómetro 	Leer señales de posición de los actuadores

Subsistema	Función	Opción 1	Opción 2	Opción 3
Procesamiento de datos	Procesar datos	Ordenador de placa reducida (SBC) con unidad de procesamiento de gráficos 	Ordenador personal 	
	Almacenar información	Memoria SD 	Disco duro externo 	Disco duro interno 
	Detectar y obtener color del semáforo	Máscara RGB 	Máscara HSV 	Máscara en HSL 
	Extraer características	YOLO V3	Faster RCNN	HOG + CSS
	Clasificar objetos		AdaBoost Decision Tree	AdaBoost Decision Stump
	Seguir Peatones	Algoritmo Camshift	Filtro de Kalman	Filtro de partículas
	Contar Peatones	BRISK	FAST	SIFT
	Generar señal de control	PID	Compensador atraso-adelanto	Deep Q-Network

Subsistema	Función	Opción 1	Opción 2	Opción 3
Energía	Almacenar energía	Batería de litio 	Batería de níquel cadmio 	Batería de plomo-ácido 
	Acondicionar energía	Convertor Step up/down 	Regulador de tensión lineal 	Regulador electromecánico 
Actuadores	Activar/desactivar el suministro de energía	Interruptor 	Botón 	Perilla 
	Accionar direccionamiento de la cámara	Motor a pasos 	Servomotor 	Motor DC 

Subsistema	Función	Opción 1	Opción 2	Opción 3
Mecánico	Albergar/ soportar componentes	Caja plástica impermeable 	Caja metálica impermeable 	Encapsulado cilíndrico 
	Fijación del sistema al poste	Abrazaderas 	Adaptador de montaje atornillado 	Garra 
	Direccionar la cámara	Plataforma plegable 	Plataforma pan tilt 	Brazo robótico 

3.5 Conceptos de solución

En la presente sección se muestran y explican los 3 conceptos de solución señalados en la matriz morfológica.

3.5.1 Concepto de solución 1

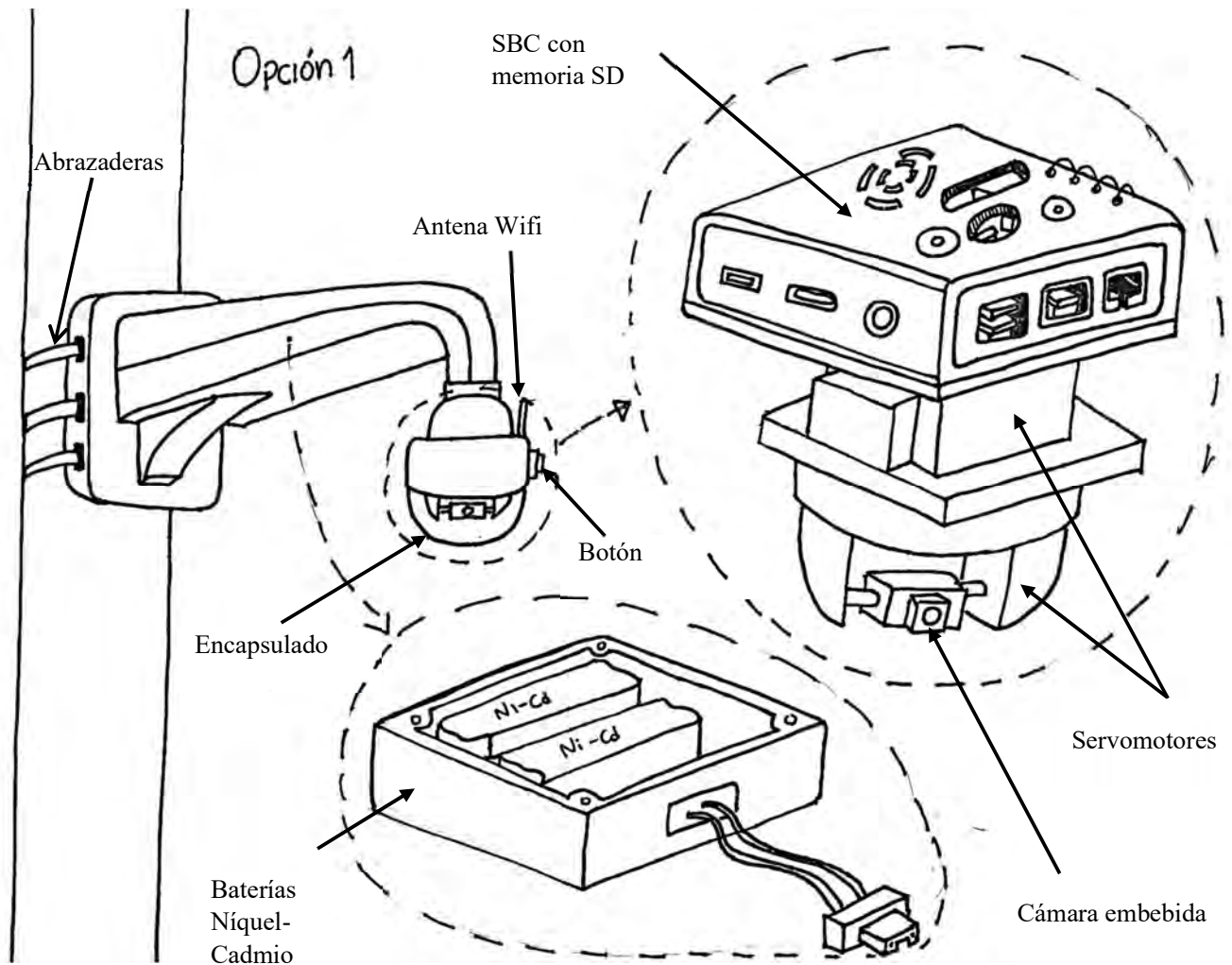


Figura 3-4 Concepto de solución 1. Elaboración propia.

El primer concepto (*Figura 3-4*) plantea el uso de una cámara embebida ubicada en una plataforma pan-tilt la cual es montada en un poste mediante el uso de abrazaderas. Los actuadores de la plataforma son servomotores, uno para cada grado de libertad, dándole al sistema la capacidad de rotar en el eje horizontal y vertical. Estos actuadores son

controlados mediante un algoritmo PID. El procesamiento de datos para el control del sistema y conteo de peatones es realizado por un ordenador de placa reducida (SBC) que cuenta con una memoria SD, en la cual se almacenarán los videos e información acerca del flujo peatonal. Tanto la plataforma como el SBC se encuentran dentro de un encapsulado el cual proveerá protección ante polvo y lluvia (IP54). En cuanto a la alimentación del sistema, esta se realizará mediante el uso de una batería de níquel cadmio, la cual se encuentra en una posición estática entre la unión del soporte con el poste y la plataforma. La energía proveniente de esta batería es acondicionada a los distintos componentes mediante el uso de convertidores Step Up/Down. El sistema comenzará a ser energizado con el uso de un botón.

Con respecto al software del sistema, se está considerando una interfaz por celular en forma de app, la cual se conectará mediante una conexión Wi-Fi, es por esta razón que este concepto de solución cuenta con una antena que pueda recepcionar las señales por dicho medio. Otro aspecto por considerar en este rubro son los algoritmos para lograr la recolección de datos acerca del flujo peatonal. En primer lugar, la detección del color del semáforo se logrará mediante una máscara en el espacio de colores HSV. En segundo lugar, la detección de peatones se realizará utilizando YOLO v3, un detector de una etapa. Luego, el seguimiento de peatones es realizado utilizando un algoritmo basado en el filtro de Kalman. Finalmente, la cantidad de peatones que pasan por la intersección en un ciclo peatonal se obtiene utilizando un algoritmo tipo BRISK.

3.5.2 Concepto de solución 2

El segundo concepto (*Figura 3-5*) se basa en el uso de una cámara digital ubicada en una plataforma pan-tilt la cual es posicionada en un poste mediante un adaptador de montaje atornillado. Los actuadores de la plataforma son dos motores DC, dándole al sistema dos grados de libertad, los cuales le permitirán rotar en el eje horizontal y vertical. La posición

de los actuadores es controlada mediante un compensador en atraso-adelanto y medida utilizando un Encoder por motor. El procesamiento de datos para el control del sistema y conteo de peatones es realizado por un ordenador de placa reducida (SBC) que está conectada a un disco duro externo, en el cual se almacenarán los videos e información acerca del flujo peatonal. Para la protección del sistema se usa una caja impermeable IP 54, en esta se encontrará el SBC, la cámara y el disco duro externo. Cabe resaltar que la caja contará con una luna transparente en el lugar donde se posicionará el lente de la cámara, de tal manera que esta pueda grabar con normalidad. En cuanto a la alimentación del sistema, este se realizará mediante el uso de una batería de litio, que se encontrará debajo de los motores dentro de una caja metálica. La energía proveniente de esta batería es acondicionada a los distintos componentes mediante el uso de reguladores de tensión lineal. El sistema comenzará a ser energizado con el uso de un interruptor.

Dentro del software del sistema, se está considerando una interfaz en computadora, que se conectará mediante una conexión Wi-Fi, es por esta razón que este concepto de solución contará con una antena que pueda recepcionar las señales por dicho medio. En esta alternativa de solución también se tienen en cuenta los algoritmos para lograr la recolección de datos acerca del flujo peatonal. La detección del color del semáforo se logra mediante una máscara en el espacio de colores RGB. Seguidamente, la detección de peatones se ejecuta utilizando HOG como extractor de características y un clasificador entrenado utilizando AdaBoost. Posteriormente, el seguimiento de peatones es realizado utilizando un filtro de partículas. Por último, la cantidad de peatones que pasan por la intersección en un ciclo peatonal es obtenida utilizando un algoritmo tipo FAST.

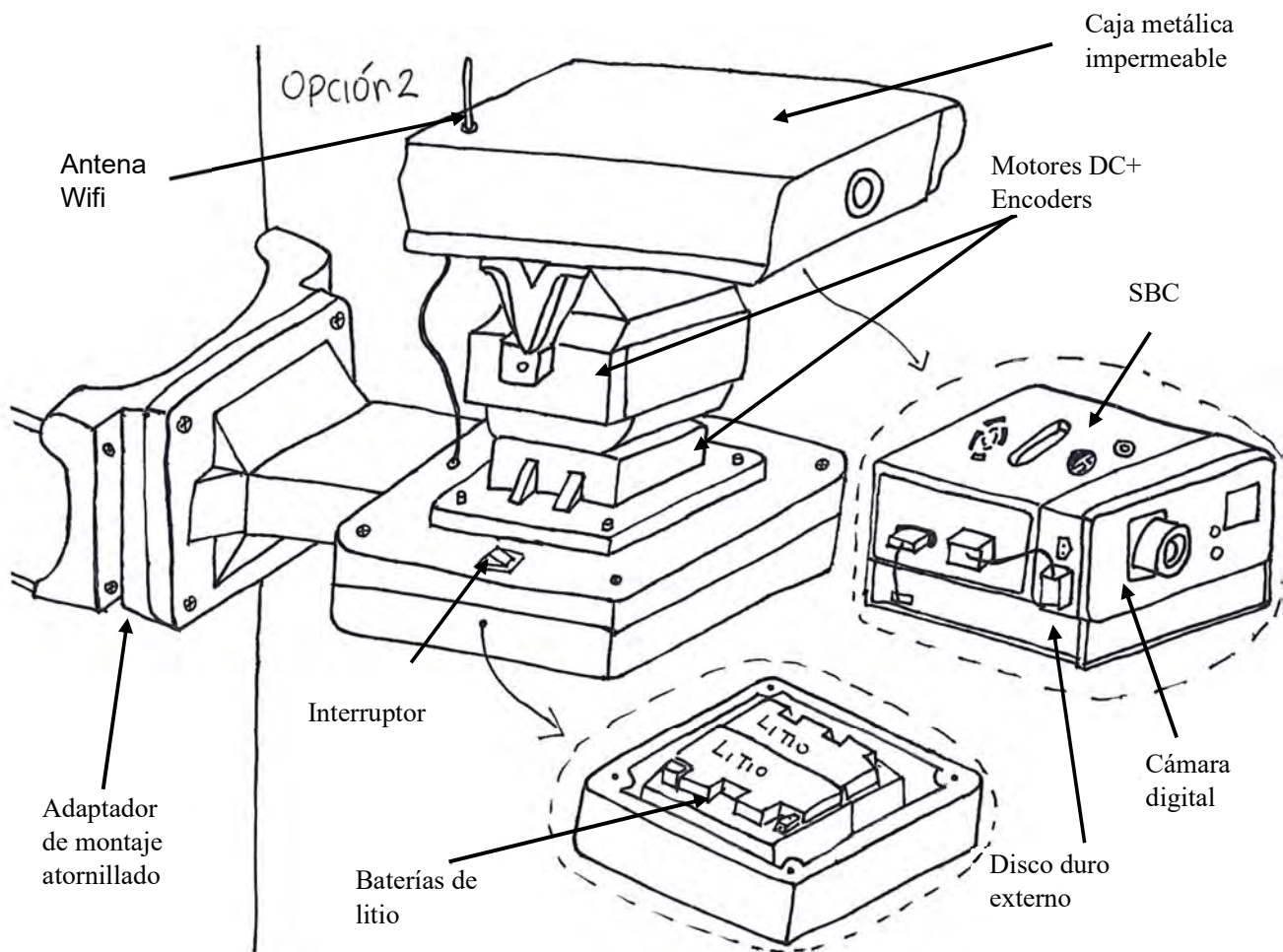


Figura 3-5 Concepto de solución 2. Elaboración propia.

3.5.3 Concepto de solución 3

El concepto de solución 3 (Figura 3-6) utiliza un brazo robótico el cual sostiene una cámara embebida que se encuentra dentro de una caja hermética de plástico IP 54. La sujeción de este sistema al poste es mediante abrazaderas. Esta alternativa cuenta con 4 grados de libertad accionados por servomotores cuya posición es controlada de manera no lineal mediante una *Deep Q Network*. Los datos acerca de la posición de los servomotores son obtenidos leyendo la señal proveniente de los potenciómetros internos. Adicionalmente, el procesamiento de datos para adquirir información del flujo peatonal y para controlar la posición del brazo robótico es realizado por una CPU de tamaño

reducido la cual almacena el video grabado y la información relacionada al conteo de peatones dentro de su disco duro interno. Con respecto a la alimentación del sistema, se activa/desactiva mediante el uso de una perilla y se utilizará una batería de plomo-ácido que estará posicionada en una caja externa que se sujetará al poste con abrazaderas.

El presente concepto de solución considera en el ámbito de software una interfaz en computadora, la cual se conectará mediante una conexión Wi-Max, por consiguiente, contará con una antena que pueda recepcionar las señales por dicho medio. Para la detección del color del semáforo se considera el uso de una máscara en el espacio de colores HSL. Además, la detección de peatones se consigue utilizando Fast RCNN y se

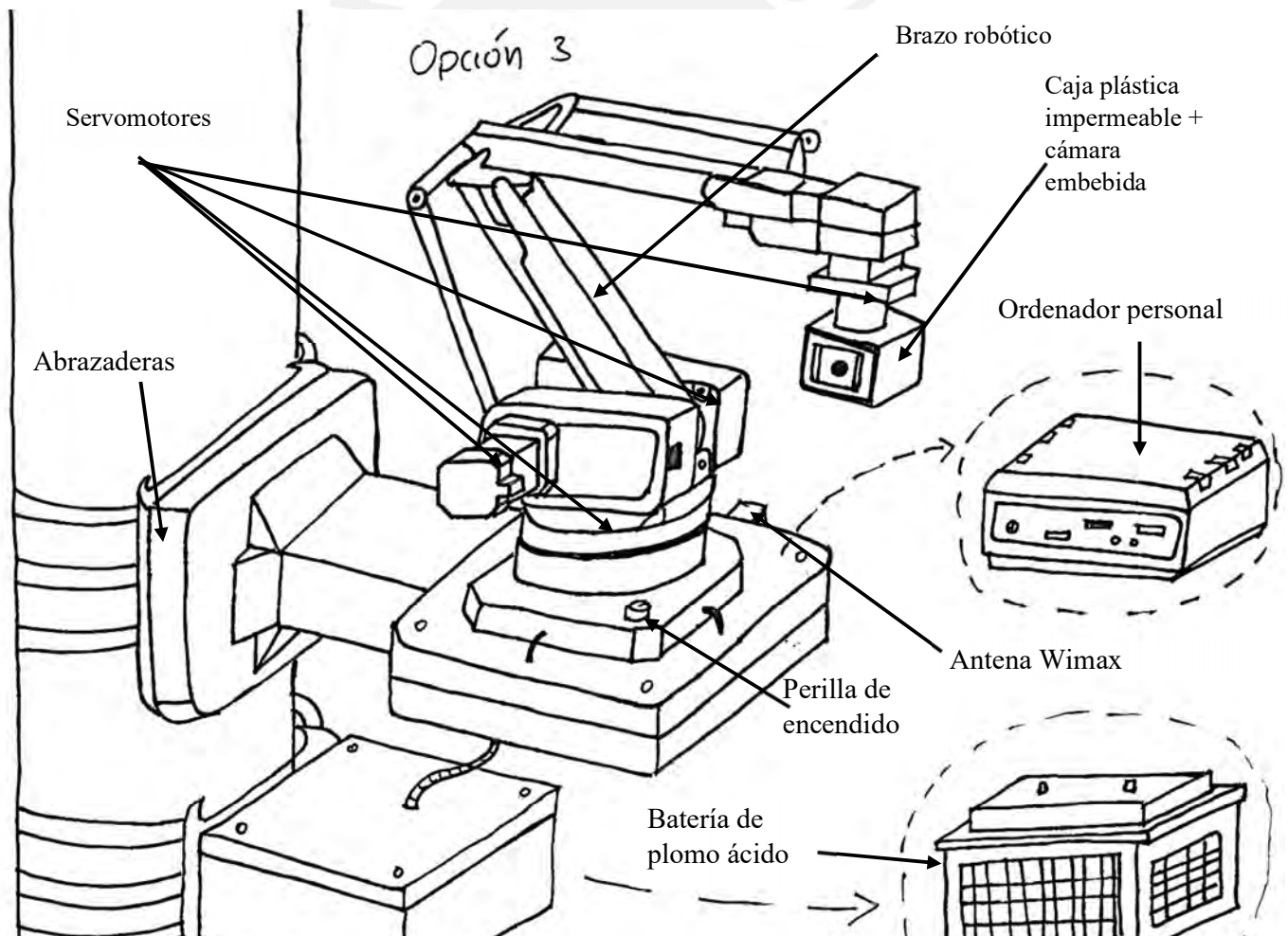


Figura 3-6 Concepto de solución 3. Elaboración propia.

entrenará el clasificador utilizando AdaBoost. Luego, el seguimiento de peatones es realizado utilizando el algoritmo Camshift. Para finalizar, el conteo se realizará mediante SIFT.

3.6 Evaluación de los conceptos de solución

Con el objetivo de identificar el concepto de solución óptimo, se evalúan las alternativas presentadas en la sección 3.5 considerando criterios técnicos y económicos. Dichos criterios tendrán un peso relativo (g) cuyo valor irá del 0 hasta el 4, el significado de cada valor se muestra en la *Tabla 3-3*.

Tabla 3-3 Pesos relativos. Elaboración propia.

Valor (g)	Significado
0	Nada importante
1	Poco importante
2	Importante
3	Muy importante
4	Bastante importante

Asimismo, cada solución tendrá un puntaje (p), el cual cuantifica qué tanto cumple con cada criterio. El significado de cada puntaje se muestra en la *Tabla 3-4*

Tabla 3-4. Escala de efectividad de las soluciones. Elaboración propia.

Valor (p)	Significado
0	No satisface
1	Mínimamente aceptable
2	Suficiente
3	Bien
4	Muy bien (Solución ideal)

3.6.1 Criterios Técnicos

En este rubro se ha decidido elegir los siguientes criterios los cuales son: Complejidad, uso de la fuerza, dimensiones, peso, rapidez de procesamiento, exactitud de conteo, exactitud de movimiento y uso de la energía. Dichos criterios son explicados en la *Tabla 3-5*.

Tabla 3-5 Criterios Técnicos. Elaboración propia.

Criterio Técnico	Descripción
Complejidad	La complejidad de un sistema está referida a la manera en la cual los componentes están distribuidos dentro del sistema, así como la cantidad de dichos elementos.
Uso de la fuerza	Este criterio incide en el torque utilizado por los motores para poder realizar los movimientos necesarios para direccionar la cámara. Es necesario conocer este aspecto para no sobre exigir a los actuadores.
Dimensiones	Un aspecto para considerar en este sistema son las dimensiones máximas debido a que estas no deben interferir con el ambiente.
Peso	Se evalúa el peso total para calcular la efectividad del elemento que sujeta el sistema al poste.
Rapidez de procesamiento	Se tiene en cuenta la rapidez del procesamiento de datos para conocer si el control de posición del sistema se puede dar a tiempo.
Exactitud de conteo	Este criterio está referido al grado de concordancia entre el número de peatones contados y número real de peatones.
Exactitud de movimiento	Este aspecto se refiere a la diferencia entre la posición real de los actuadores y la posición deseada para lograr el direccionamiento óptimo de la cámara.
Uso de la energía	Es referido al uso de energía por los componentes. Importante al momento de considerar las dimensiones de la batería a utilizar y su tiempo de uso.

3.6.2 Criterios Económicos

En este rubro se ha decidido elegir los siguientes criterios los cuales son: Costo de la tecnología, número de piezas, disponibilidad de los componentes y desperdicios. Dichos criterios son explicados en la *Tabla 3-6*

Tabla 3-6 . Criterios Económicos. Elaboración propia.

Criterio Económico	Descripción
Costo de la tecnología	Se evaluará el costo de los componentes de cada solución. Este costo posiblemente está relacionado con la complejidad tecnológica del sistema
Número de piezas	Se tendrá en consideración la cantidad de componentes del sistema debido a que estos pueden disminuir o aumentar los costos de fabricación y mantenimiento.
Disponibilidad de los componentes	Este criterio incide en la facilidad de encontrar los componentes de cada concepto de solución en el mercado local.
Facilidad de mantenimiento	Este criterio evalúa disponibilidad de personal capacitado para ejecutar mantenimiento a los componentes del sistema debido a que los costos de una máquina pueden estar estrechamente relacionados a su mantenimiento.
Eficiencia energética	Se considerará las pérdidas de energía en forma de energía calorífica, lumínica y sonora.

3.6.3 Obtención del concepto óptimo

Teniendo en cuenta los criterios explicados en las secciones 3.6.1 y 3.6.2, en la *Tabla 3-7* se detallarán los puntajes asignados, así como los promedios de cada concepto de solución con respecto al puntaje ideal.

Tabla 3-7. Evaluación técnica-económica. Elaboración propia.

TÉCNICOS		Solución 1			Solución 2		Solución 3		Ideal	
Nro.	Criterio	g	p	pxg	p	pxg	p	pxg	p	Pxg
1	Complejidad	2	4	8	4	8	2	4	4	8
2	Uso de la fuerza	4	4	16	3	12	3	12	4	16
3	Dimensiones	3	3	9	2	6	2	6	4	12
4	Peso	3	3	9	3	9	2	6	4	12
5	Rapidez de procesamiento	3	3	9	2	6	3	9	4	12
6	Precisión de conteo	4	3	12	3	12	4	16	4	16
7	Precisión de movimiento	2	2	4	2	4	3	6	4	8
8	Uso de la energía	2	3	6	3	6	2	4	4	8
Suma			25	73	22	63	21	63	32	92
Promedio			0.78	0.79	0.69	0.68	0.66	0.68	1	1
ECONÓMICOS		Solución 1			Solución 2		Solución 3		Ideal	
Nro.	Criterio	g	p	pxg	p	pxg	p	pxg	p	pxg
1	Costo de la tecnología	3	3	9	2	6	2	6	4	12
2	Número de piezas	2	3	6	3	6	2	4	4	8
3	Disponibilidad de los componentes	3	3	9	4	12	3	9	4	12
4	Facilidad de mantenimiento	4	3	12	3	12	2	8	4	16
5	Eficiencia energética	2	3	6	4	8	3	6	4	8
Suma			15	42	16	44	12	33	20	56
Promedio			0.75	0.75	0.80	0.79	0.60	0.59	1	1

Los promedios obtenidos en la *Tabla 3-7* son plasmados en un gráfico de dispersión como se puede observar en la *Figura 3-7*.

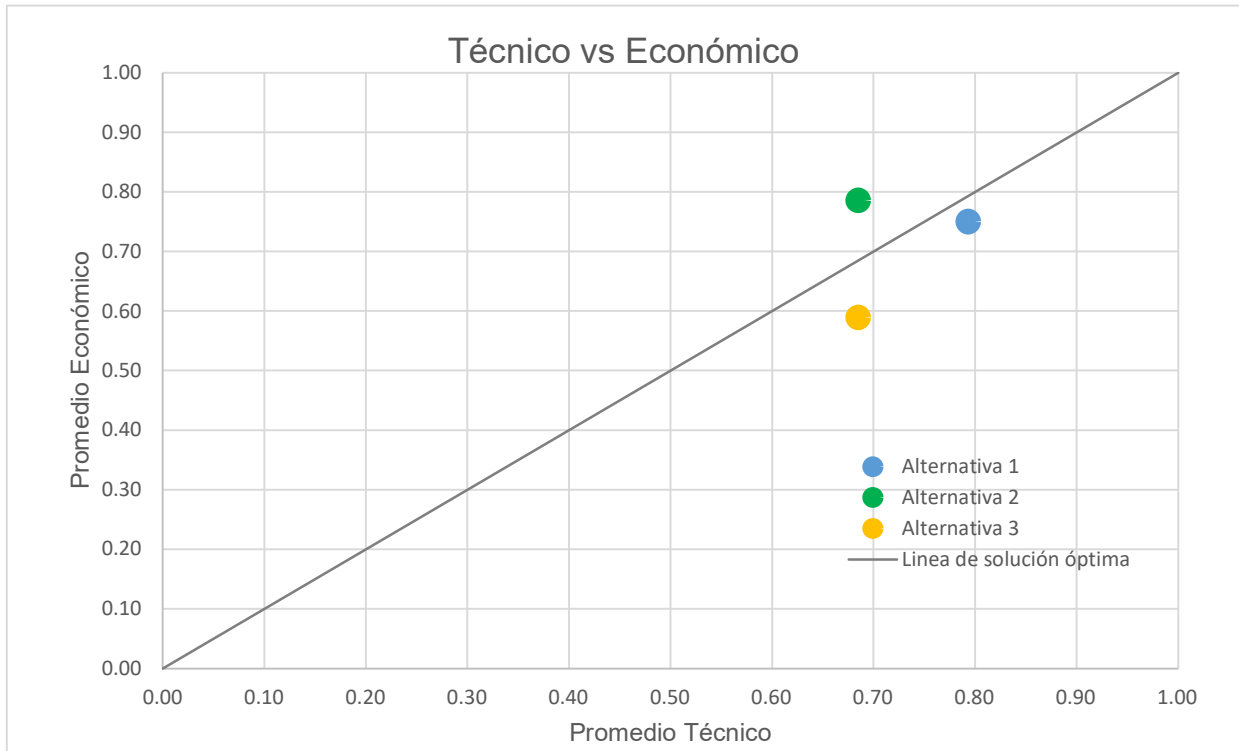


Figura 3-7 Gráfico técnico-económico entre las soluciones. Elaboración propia.

Analizando la *Figura 3-7* podemos concluir que la alternativa 1 es la adecuada debido a que es la que tiene la menor distancia a la línea de solución óptima. Además, se puede ver que las demás alternativas están alejadas de manera más considerable con respecto a la recta ideal a comparación de la primera opción. También es oportuno subrayar que la alternativa elegida tiene en promedio mayor puntaje que las demás, reafirmando así su conveniencia.

En el capítulo 3 se desarrolló el diseño de la propuesta de solución a través de un proceso que involucró determinar la lista de exigencias las cuales serían utilizadas para definir las funciones necesarias para cumplir con dichos requerimientos. Luego se buscaron principios de solución para realizar estas funciones. Finalmente se plantearon conceptos de solución y mediante una evaluación económica-técnica, se eligió el concepto óptimo.

Conclusiones

- Se consiguió proponer un concepto de solución óptimo para la problemática de la falta de automatización en la toma de datos de flujo peatonal. Dicho concepto de solución es el más adecuado en comparación a los demás conceptos propuestos debido a factores técnicos y económicos. Por el lado técnico, esta alternativa puede cumplir con la función principal de una manera igual de eficiente en comparación a las otras alternativas utilizando menor complejidad mecánica y peso. Con respecto al lado económico, la propuesta elegida cuenta con componentes y algoritmos de fácil obtención, así como con un número de piezas reducido en comparación a las demás alternativas.
- En el presente trabajo académico se han analizado las diferentes tecnologías relacionadas al conteo peatonal y la automatización de plataformas para cámaras. Gracias a esta investigación se puede afirmar que la tecnología necesaria para realizar el concepto de solución óptimo está disponible, pero la integración entre las diversas tecnologías carece de extensa documentación debido a que es un campo emergente.
- Los conteos de peatones son cada vez más necesarios para el planeamiento y verificación de políticas públicas relacionadas con el tráfico peatonal; sin embargo, el proceso manual es costoso debido a los pagos a la mano de obra y las cuentas obtenidas no tienen manera de ser verificadas. Teniendo esto en cuenta, se puede afirmar que la automatización del proceso es relevante actualmente debido a que potencialmente solucionaría los mayores problemas del conteo manual y brindaría información necesaria para tomar decisiones como evaluar la construcción de intersecciones diagonales u optimizar los ciclos semafóricos.

Bibliografía

- Alm, C.-A., Lundberg, S., Sagefalk, & Willy. (2009). *Estados Unidos Patent No. US9007474B2*.
- Benenson, R., Omran, M., Hosang, J., & Schiele, B. (2014). Ten Years of Pedestrian Detection,. *European Conference on Computer Vision* (pp. 613–627). Zurich: Springer.
- Boarnet, M. G. (1997). Infrastructure services and the productivity of public capital: The case of streets and highways. *National Tax Journal*, 50(1), 39–57.
- Cardenas, L., León, J., Pandura, J., Quispe, E., & Ramirez, I. (2018, Diciembre 9). *Análisis, diagnóstica y diseño de la intersección vial: Av. Javier Prado con Av. Aviación*. Retrieved from https://issuu.com/estefaniequispesalas/docs/urba_final_brochure_pdf
- Chen, G., Ding, Y., Xiao, J., & Han, T. (2013). Detection Evolution with Multi-order Contextual Co-occurrence. *2013 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1798-1805). Portland: IEEE.
- Deregibus, B., Juan, E., Terraza, H., Soulier, M., Ramirez, I., Schwint, A., & Moscoso, G. (2016, Julio). *BID-Publicaciones*. Retrieved from Voces emergentes: Percepciones sobre la calidad de vida urbana en América Latina y el Caribe: <https://publications.iadb.org/es/voces-emergentes-percepciones-sobre-la-calidad-de-vida-urbana-en-america-latina-y-el-caribe>
- Doig, J. (2011). *Análisis del nivel de servicio peatonal en la ciudad de Lima*. Lima: Repositorio de Tesis PUCP.

- Dong, E., Yan, S., & Wei, K. (2015). Design and Implementation of a Moving Object tracking system. *2015 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 274-279). Beijing: IEEE.
- Fernandes, F., Weigel, L., Jung, C., & Navaux, P. (2016). Evaluation of Histogram of Oriented Gradients Soft Errors Criticality. *ACM Transactions on Architecture and Code Optimization (TACO)*, 38:1-38:25.
- Fernández-Garza, A., & Hernández-Vega, H. (2019). Estudio de la movilidad peatonal en un centro urbano: un caso en Costa Rica. *Revista Geográfica de América Central. Número 62*, 267-300.
- Gaddigoudar, P., Balihalli, T., Ijantkar, S., Iyer, N., & Maralappanavar, S. (2017). Pedestrian detection and tracking using particle filtering . *2017 International Conference on Computing, Communication and Automation (ICCCA)* (pp. 110-115). Greater Noida: IEEE.
- Huang, S., Chen, W., Yu, R., Yang, X., & Dong, D. (2018). Predicting Pedestrian Counts for Crossing Scenario Based on. *Journal of Advanced Transportation*, 1-11.
- Hymel, K. (2009). Does traffic congestion reduce employment growth? *Journal of Urban Economics*, 65(2), 127–135.
- Instituto de Opinión Pública-PUCP. (2018, Marzo 21). *Encuesta Lima cómo vamos*. Retrieved from VIII Informe de percepción de percepción sobre la calidad de vida: http://www.limacomovamos.org/cm/wp-content/uploads/2018/03/EncuestaLimaC%C3%B3moVamos_2017.pdf

- Jakhar, D., & Kaur, I. (2019). *Artificial intelligence, machine learning and deep learning: definitions and differences*. Oxford: Blackwell Scientific Publications.
- Jeong, C., & Choi, S. (2016). A Comparison of Keypoint Detectors in the Context. *2016 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 1179-1181). Jeju: IEEE.
- Jeong, C., Kim, M., & Shin, H. (2018). Soft assignment and multiple keypoint analysis-based pedestrian counting method. *2018 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 1479-1481). Jeju Island: The Korean Institute of Communications and Information Sciences. Retrieved from Soft assignment and multiple keypoint analysis-based pedestrian counting method: <https://ieeexplore-ieee.org.ezproxybib.pucp.edu.pe/document/8539499>
- Jimenez, J., Martin, A., Uc, V., & Espinosa, A. (2017). Mexican Sign Language Alphanumerical Gestures Recognition using 3D Haar-like Features. *IEEE Latin America Transactions*, 2000-2005.
- Kim, D., Kim, K., & Sungjoo, P. (2019). Automatic PTZ Camera Control Based on Deep-Q. *2019 International Conference on Electronics, Information, and Communication (ICEIC)*. Auckland: IEEE. Retrieved from IEEE Xplore.
- Lee, S., & Zixiang, X. (2015). A real-time face tracking system based on a single PTZ camera. *2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)* (pp. 568-572). Chengdu: IEEE.
- Lindsey, G., Nordback, K., & Figliozzi, M. A. (2016). Institutionalizing Bicycle and Pedestrian. *93rd Annual Meeting of the Transportation Research* (pp. 134-142). Washington DC: Sage Journals.

- Municipalidad de Miraflores. (2017). *Plan de Movilidad Urbana de Miraflores*. Retrieved from http://www.miraflores.gob.pe/Gestorw3b/files/pdf/10773-29116-plan_de_movilidad_web.pdf
- Municipalidad de San Isidro. (2016). *Municipalidad de San Isidro*. Retrieved from Plan de Movilidad Urbana del Distrito de San Isidro: <http://msi.gob.pe/portal/plan-de-movilidad-urbana/>
- Orellana, D., Hermida, C., & Osorio, P. (2017). Comprendiendo los patrones de movilidad de ciclistas y peatones. Una síntesis de literatura. *Transporte y Territorio/16*, 108. Retrieved from <http://revistascientificas.filo.uba.ar/index.php/rtt/article/view/3608/3303>
- Ouyang, W., & Wang, X. (2013). Single-Pedestrian Detection Aided by Multi-pedestrian Detection. *2013 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3198-3205). Portland: IEEE.
- Pahl, G., Beitz, W., Feldhusen, J., & Grote, K. (2007). *Engineering Design*. Londres: Springer.
- Peña, C. (2019, Julio 26). Miraflores inaugura primer cruce peatonal múltiple. (L. Vargas, Interviewer) Retrieved from <https://panamericana.pe/buenosdiasperu/locales/270803-miraflores-inauguran-primer-cruce-peatonal-multiple>
- Quiong, R. (2015). Multi-pedestrian tracking for far-infrared pedestrian detection on-board using particle filter. *2015 IEEE International Conference on Imaging Systems and Techniques (IST)*. Macao: IEEE.

- Rauf, R., Shahid, S., Ziauddin, S., & Safi, A. (2017). Pedestrian detection using HOG, LUV and optical flow as features with AdaBoost as classifier. *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)* (pp. 1-4). Oulu: IEEE.
- Sung, C., & Rus, D. (2015). Foldable Joints for Foldable Robots. *Journal of Mechanisms and Robotics*, 1-13.
- Tettamanti, T., Csikòs, A., Kis, K., Viharos, Z., & Varga, I. (2017, Septiembre 4). *Academic Search Complete*. Retrieved from PATTERN RECOGNITION BASED SPEED FORECASTING METHODOLOGY FOR URBAN TRAFFIC NETWORK.:
<http://web.b.ebscohost.com.ezproxybib.pucp.edu.pe:2048/ehost/detail/detail?vid=3&sid=6f416896-4bc5-4c17-883e-f22dc4073862%40pdc-v-sessmgr03&bdata=Jmxhbmc9ZXMmc2l0ZT1laG9zdC1saXZl#db=a9h&AN=133490388>
- TomTom. (2018). *TomTom*. Retrieved from Traffic Index:
https://www.tomtom.com/en_gb/traffic-index/ranking
- U.S Department of Transportation. (2016, Octubre). *Policy and Governmental Affairs. Office of Highway Policy Information*. Retrieved from Traffic Monitoring Guide: <https://www.fhwa.dot.gov/policyinformation/tmguide/>
- Uddin, M., & Akhi, A. (2016). Horse Detection Using Haar Like Features. *International Journal of Computer Theory and Engineering, Vol. 8, No. 5*, 415-418.
- Ukida, H. (2010). Object Tracking System by Pan-Tilt Moving. *SICE Annual Conference 2010* (pp. 99-104). Taiwan: IEEE.

- Varthakavi, S. S., Babu, D., & Reddy, L. (2019). Analysis of Preprocessing Algorithms for Face Detection using KNN and SVM Classifiers. *Grenze International Journal of Engineering & Technology (GIJET)*. 2019, Vol. 5 Issue , 32-39.
- Wang, X., Han, T. X., & Yan, S. (2009). An HOG-LBP Human Detector with Partial Occlusion Handling. *2009 IEEE 12th International Conference on Computer Vision* (pp. 32-39). Kyoto: IEEE.
- Yan, J., Zhan, X., Lei, Z., & Lin, S. (2013). Robust Multi-resolution Pedestrian Detection in Traffic Scenes. *2013 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3033-3040). Portland: IEEE.
- Yang, H., Kaan, O., & Bartin, B. (2010). Investigating the performance of automatic counting sensors for pedestrian traffic data collection. *Proceedings of the 12th World Conference on Transport* (pp. 1-11). Portugal: WCTRS.
- Yosafat, R., Machbub, C., & Hidayat, E. (2017). Design and Implementation of Pan-Tilt. *2017 7th IEEE International Conference on System Engineering and Technology (ICSET)* (pp. 217-222). Shah Alam: IEEE.
- Yu, Z., Gong, C., Yang, L., & Li, B. (2014). PEDESTRIAN COUNTING BASED ON SPATIAL AND TEMPORAL ANALYSIS. *2014 IEEE International Conference on Image Processing (ICIP)* (pp. 2432-2436). Paris: IEEE.
- Zeng, X., Ouyang, W., & Xiaogang, W. (2013). Multi-stage Contextual Deep Learning for Pedestrian Detection. *2013 IEEE International Conference on Computer Vision* (pp. 121-128). Sydney: IEEE.

Zhang, S., Bauckhage, C., & Cremers, A. (2014). Informed Haar-Like Features Improve Pedestrian Detection. *2014 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 947-954). Columbus: IEEE.

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems & Technology*, 5(3), 222–23.

