



Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

**DISEÑO DE INVESTIGACIÓN DE UN PROTOTIPO PARA EL ANÁLISIS Y CLASIFICACIÓN  
DE INFORMACIÓN VISUAL PARA LA AUTOMATIZACIÓN DEL MONITOREO DE CÁMARAS  
IP CONECTADAS A UNA RED LOCAL EN UNA VIVIENDA DE LA CIUDAD DE GUATEMALA**

**Julia Argentina Sierra Herrera**

Asesorado por M. A. Isiris Yassmín Córdova Grajeda

Guatemala, agosto de 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DISEÑO DE INVESTIGACIÓN DE UN PROTOTIPO PARA EL ANÁLISIS Y CLASIFICACIÓN DE INFORMACIÓN VISUAL PARA LA AUTOMATIZACIÓN DEL MONITOREO DE CAMARAS IP CONECTADAS A UNA RED LOCAL EN UNA VIVIENDA DE LA CIUDAD DE GUATEMALA**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA  
FACULTAD DE INGENIERÍA  
POR

**JULIA ARGENTINA SIERRA HERRERA**

ASESORADO POR M. A. ISIRIS YASSMÍN CÓRDOVA GRAJEDA

AL CONFERÍRSELE EL TÍTULO DE

**INGENIERA EN CIENCIAS Y SISTEMAS**

GUATEMALA, AGOSTO DE 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA  
FACULTAD DE INGENIERÍA



**NÓMINA DE JUNTA DIRECTIVA**

DECANA	Inga. Aurelia Anabela Cordova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Kevin Vladimir Cruz Lorente
VOCAL V	Br. Fernando José Paz González
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

**TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO**

DECANA	Inga. Aurelia Anabela Cordova Estrada
EXAMINADOR	Ing. César Augusto Fernández Cáceres
EXAMINADOR	Ing. César Rolando Batz Saquimux
EXAMINADOR	Ing. Herman Igor Véliz Linares
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

## **HONORABLE TRIBUNAL EXAMINADOR**

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**DISEÑO DE INVESTIGACIÓN DE UN PROTOTIPO PARA EL ANÁLISIS Y CLASIFICACIÓN DE INFORMACIÓN VISUAL PARA LA AUTOMATIZACIÓN DEL MONITOREO DE CAMARAS IP CONECTADAS A UNA RED LOCAL EN UNA VIVIENDA DE LA CIUDAD DE GUATEMALA**

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Postgrado con fecha 28 de mayo de 2022.



**Julia Argentina Sierra Herrera**



**EEPFI-PP-0712-2022**

Guatemala, 28 de mayo de 2022

**Director**  
**Carlos Gustavo Alonzo**  
**Escuela De Ingenieria En Sistemas**  
**Presente.**

**Estimado Ing. Alonzo**

Reciba un cordial saludo de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería.

El propósito de la presente es para informarle que se ha revisado y aprobado el Diseño de Investigación titulado: **PROTOTIPO PARA EL ANALISIS Y CLASIFICACION DE INFORMACION VISUAL PARA LA AUTOMATIZACION DEL MONITOREO DE CAMARAS IP CONECTADAS A UNA RED LOCAL EN UNA VIVIENDA DE LA CIUDAD DE GUATEMALA**, el cual se enmarca en la línea de investigación: **Área de Innovación - Dispositivos y sistemas para el desarrollo de sistemas inteligentes.**, presentado por la estudiante **Julia Argentina Sierra Herrera** carné número **201503865**, quien optó por la modalidad del "PROCESO DE GRADUACIÓN DE LOS ESTUDIANTES DE LA FACULTAD DE INGENIERÍA OPCIÓN ESTUDIOS DE POSTGRADO". Previo a culminar sus estudios en la Maestría en ARTES en Tecnologías De La Inf. Y La Comunicacion.

Y habiendo cumplido y aprobado con los requisitos establecidos en el normativo de este Proceso de Graduación en el Punto 6.2, aprobado por la Junta Directiva de la Facultad de Ingeniería en el Punto Décimo, Inciso 10.2 del Acta 28-2011 de fecha 19 de septiembre de 2011, firmo y sello la presente para el trámite correspondiente de graduación de Pregrado.

Atentamente,

*"Id y Enseñad a Todos"*

  
Isiris Y. Córdova Grajeda  
INGENIERA EN SISTEMAS Y SISTEMAS  
Colegiado No.13783

Mtra. Isiris Yassmín Córdova Grajeda  
Asesor(a)

  
Mtro. Marlon Arístido Pérez Turk  
Coordinador(a) de Maestría





Mtro. Edgar Darío Álvarez Cotí  
Director  
Escuela de Estudios de Postgrado  
Facultad de Ingeniería





EPP-EICS-0712-2022

El Director de la Escuela De Ingenieria En Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del Asesor, el visto bueno del Coordinador y Director de la Escuela de Estudios de Postgrado, del Diseño de Investigación en la modalidad Estudios de Pregrado y Postgrado titulado: **PROTOTIPO PARA EL ANALISIS Y CLASIFICACION DE INFORMACION VISUAL PARA LA AUTOMATIZACION DEL MONITOREO DE CAMARAS IP CONECTADAS A UNA RED LOCAL EN UNA VIVIENDA DE LA CIUDAD DE GUATEMALA**, presentado por el estudiante universitario **Julia Argentina Sierra Herrera**, procedo con el Aval del mismo, ya que cumple con los requisitos normados por la Facultad de Ingeniería en esta modalidad.

ID Y ENSEÑAD A TODOS

Ing. Carlos Gustavo Alonzo  
Director  
Escuela De Ingenieria En Sistemas

Guatemala, mayo de 2022

Decanato  
Facultad de Ingeniería  
24189101- 24189102  
secretariadecanato@ingenieria.usac.edu.gt

LNG.DECANATO.OI.589.2022

La Decana de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al Trabajo de Graduación titulado: **DISEÑO DE INVESTIGACIÓN DE UN PROTOTIPO PARA EL ANÁLISIS Y CLASIFICACIÓN DE INFORMACIÓN VISUAL PARA LA AUTOMATIZACIÓN DEL MONITOREO DE CÁMARAS IP CONECTADAS A UNA RED LOCAL EN UNA VIVIENDA DE LA CIUDAD DE GUATEMALA**, presentado por: **Julia Argentina Sierra Herrera**, después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:



Inga. Aurelia Anabela Cordova Estrada

Decana

Guatemala, agosto de 2022

AACE/gaoc

## **ACTO QUE DEDICO A:**

**Dios**

Por darme la sabiduría que necesito para cumplir con mis metas propuestas.

**Mis padres**

Thelma Herrera y Guillermo Sierra (q. e. p. d.), por su amor, cariño y comprensión en todos los momentos de mi vida, por cada uno de sus sacrificios para ayudarme a cumplir mis metas.

**Mis hermanas**

Thelma y Martha Sierra, por darme su apoyo incondicional.



## **AGRADECIMIENTOS A:**

**Universidad de San  
Carlos de  
Guatemala**

Por ser la casa de estudio que me permitió desarrollarme de manera académica, a través del conocimiento brindado y experiencia.

**Mi padre**

Guillermo Sierra (q. e. p. d.), por criarme con amor y ayudarme en cada una de mis metas de manera incondicional.

**Mi madre**

Thelma Herrera, por acompañarme en cada uno de los momentos de mi vida y darme las fuerzas para continuar día a día con mis metas.

**Mis hermanas**

Martha y Thelma Sierra, por acompañarme en el trayecto de mi carrera e incentivarme a siempre dar lo mejor de mí.

**Mi mejor amigo**

Julio Arango, por brindarme su apoyo, por ser parte importante en mi carrera, ofreciéndome su amistad en todo momento.



## ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES.....	V
LISTA DE SÍMBOLOS .....	VII
GLOSARIO .....	IX
RESUMEN.....	XI
1. INTRODUCCIÓN .....	1
2. ANTECEDENTES .....	5
3. PLANTEAMIENTO DEL PROBLEMA .....	9
4. JUSTIFICACIÓN .....	13
5. OBJETIVOS .....	15
5.1. General.....	15
5.2. Específicos .....	15
6. NECESIDADES A CUBRIR Y ESQUEMA DE LA SOLUCIÓN .....	17
6.1. Necesidades a cubrir .....	17
6.2. Esquema de solución .....	17
7. MARCO TEÓRICO.....	21
7.1. Sistema de videovigilancia .....	21
7.1.1. Definición .....	21
7.1.2. Videovigilancia IP .....	22

7.1.3.	Arquitectura .....	23
7.1.4.	Cámara IP .....	24
7.1.5.	Transmisión de datos .....	25
7.1.6.	Monitoreo de sistemas de vigilancia.....	26
7.2.	<i>Deep learning</i> .....	27
7.2.1.	Definición.....	27
7.2.2.	Componentes .....	29
7.3.	Redes neuronales .....	31
7.3.1.	Definición.....	31
7.3.2.	Tipos de aprendizaje .....	32
7.3.2.1.	Aprendizaje supervisado .....	32
7.3.2.2.	Aprendizaje no supervisado .....	33
7.3.3.	Aplicaciones .....	33
7.3.3.1.	Concepto de procesamiento de imágenes.....	33
7.3.4.	Redes neuronales convolucionales.....	34
7.3.4.1.	Definición.....	34
7.3.4.2.	Detección de bordes en imágenes .....	35
7.3.5.	Redes neuronales recurrentes .....	35
7.3.5.1.	Definición.....	35
7.3.5.2.	Memoria a corto plazo (LSTM) .....	36
8.	PROPUESTA DE ÍNDICE DE CONTENIDO .....	37
9.	METODOLOGÍA .....	39
9.1.	Diseño de investigación .....	39
9.2.	Variables de estudio.....	39
9.3.	Alcances.....	40
9.4.	Procedimiento metodológico .....	43

9.4.1.	Observación documental .....	43
9.4.2.	Investigación documental .....	43
9.4.3.	Decisión con base a conocimientos.....	44
9.4.4.	Implementación de algoritmos de reconocimiento..	44
9.4.5.	Implementación de prototipo.....	45
9.4.6.	Análisis de resultados.....	45
9.5.	Técnicas de recolección de información.....	46
10.	TÉCNICA DE ANÁLISIS DE INFORMACIÓN .....	47
11.	CRONOGRAMA.....	49
12.	FACTIBILIDAD DEL ESTUDIO .....	51
12.1.	Factibilidad operativa.....	51
12.2.	Factibilidad técnica .....	53
12.3.	Factibilidad económica .....	54
13.	REFERENCIAS.....	57



## ÍNDICE DE ILUSTRACIONES

### FIGURAS

1.	Arquitectura general del sistema.....	18
2.	Esquema básico de un sistema de Videovigilancia IP .....	23
3.	Esquema básico de una cámara de red.....	24
4.	Relación entre inteligencia artificial, <i>machine learning</i> y <i>deep learning</i> .....	27
5.	Diferencias entre <i>machine learning</i> y <i>deep learning</i> .....	28
6.	Modelo <i>deep learning</i> con sus diferentes etapas .....	30
7.	Estructura jerárquica de un sistema basado en RNA.....	32
8.	Cronograma .....	49

### TABLAS

I.	Variables de estudio.....	40
II.	Presupuesto .....	55



## LISTA DE SÍMBOLOS

<b>Símbolo</b>	<b>Significado</b>
%	Porcentaje
Q	Quetzales



## GLOSARIO

<b>CCTV</b>	Circuito Cerrado de Televisión. Es un sistema de transmisión y visualización de imágenes en movimiento.
<b>Inteligencia artificial</b>	Inteligencia diseñada para simular el razonamiento humano en un ámbito computacional, de manera que una maquina puede aprender, tomar decisiones con el fin de lograr resolver un problema.
<b>IP</b>	Dirección del Protocolo de Internet, este protocolo es un conjunto de reglas para la comunicación a través de internet, ya sea el envío de correo electrónico, transmisión de video o conexión a un sitio web
<b>LAN</b>	Es una red de computadoras que abarca un área reducida a una casa, un departamento o edificio.
<b><i>Router</i></b>	Es un dispositivo que ofrece una conexión Wi-Fi, que normalmente está conectado a un modem y que envía información de internet a tus dispositivos personales, como ordenadores, teléfonos, etc.



## RESUMEN

El presente trabajo plantea el diseño de un prototipo para el análisis y clasificación de información visual para la automatización del monitoreo de cámaras IP conectadas a una red local para su uso en viviendas de la Ciudad de Guatemala.

Para abordar el problema se trabaja en fases las cuales son: observación documental de procesos existentes que se utilizan para el monitoreo de manera operativa, investigación documental de métodos existentes para análisis y reconocimiento de imágenes, decisión en base a los conocimientos donde se reconoce que métodos utilizar y equipo físico, implementación de algoritmos de reconocimiento, se iniciara con el uso de los diferentes métodos para luego implementarlo en el prototipo desarrollado, para concluir con el análisis de resultados que determinara la eficiencia de los métodos utilizados.

Con esta propuesta de diseño lo que se quiere lograr es una mejora en los procesos de monitoreo actuales, así como simplificar la compatibilidad con los equipos utilizados para realizar monitoreo en viviendas de la ciudad de Guatemala.



# 1. INTRODUCCIÓN

El monitoreo de cámaras de video es un proceso que requiere de atención en tiempo completo para lograr la clasificación y análisis de la información visual, este es un proceso que aún se desarrolla de forma operativa en varios sectores de la ciudad de Guatemala, incluso existen cámaras que no son monitoreadas y solo permiten su grabación para posteriormente realizar análisis si fuera necesario, con los índices de delincuencia que reporta el país el monitoreo se convierte en una necesidad para algunos sectores con el fin de lograr asegurar bienes, recaudar evidencias y prevenir incidentes como robos entre otros.

Los sistemas de información visual como lo es el de videovigilancia requieren de automatización en sus tareas para analizar y clasificar la información para lograr el monitoreo de manera eficiente e inmediata. Los sistemas de inteligencia artificial como lo son las redes neuronales son implementados para diferentes usos como el reconocimiento de imágenes. El uso de estos algoritmos reduce el costo operativo para realizar el monitoreo; permite que el análisis y clasificación de información se realice de forma inmediata.

La mayoría de los sistemas de videovigilancia instalados en diferentes localidades de la ciudad de Guatemala no poseen un sistema que automatice las fases de análisis y clasificación de información, esto provoca que las operaciones del monitoreo en un sistema de videovigilancia se retrasen y no den respuestas de manera inmediata. Como consecuencia se reduce la eficiencia del sistema y limita a los usuarios finales a quedarse con grabaciones que se deben ser analizadas y clasificadas posteriormente con la revisión total de las grabaciones.

Actualmente los sistemas de videovigilancia estructurados por cámaras IP independientes, no cuentan con infraestructura ni sistema tecnológico que permita realizar el análisis y clasificación de la información para lograr recibir respuestas inmediatas de los sucesos que acontecen en las imágenes obtenidas del sistema de vigilancia.

El presente trabajo presenta una solución para lograr el análisis y clasificación de información visual con el fin de poder optimizar recursos operativos y tecnológicos para beneficiar a los usuarios que cuentan con sistemas de monitoreo, pero actualmente son costos en tiempo para los usuarios o requieren de recurso humano para realizarlo.

Para la ejecución de este trabajo se definen los siguientes capítulos que permitirán la investigación, diseño y desarrollo del prototipo:

Capítulo uno: presenta los trabajos realizados acerca de reconocimiento de imágenes, implementaciones para prevenir sucesos inusuales en imágenes, las diferentes aplicaciones que se han logrado con la utilización del reconocimiento de imagen.

Capítulo dos: se presenta la problemática que se resolverá, donde se da a conocer la situación actual de los índices de delincuencia en la ciudad de Guatemala, porque la población considera necesarios los sistemas de videovigilancia y qué problemas surgen al no tener automatización del sistema.

Capítulo tres: se presenta la definición del objetivo general y de los específicos que resolverán la falta de automatización, el análisis y clasificación de la información en el monitoreo de cámara IP que se presentan en el capítulo dos.

Capítulo cuatro: en este se presenta la justificación, se define porqué es importante realizar una automatización de análisis y clasificación de información implementando reconocimiento de imágenes, dando a conocer los beneficios que se obtienen al lograr esta automatización en el monitoreo y cómo estos pueden beneficiar a la población de la ciudad de Guatemala.

Capítulo cinco: en este se presentan las necesidades que debe de satisfacer el sistema que se está desarrollando, así como un esquema donde se especifica de qué manera se realizará el prototipo y cuáles son los componentes que se utilizaran para su implementación.

Capítulo seis: se define el alcance del estudio compuesto por la perspectiva investigativa, técnica y de resultados. En la cual se define lo que abarcará la implementación del prototipo y cuáles son los requisitos para lograr los objetivos del estudio, así como cuáles son los sistemas que deben ser desarrollados y de qué manera se requiere que sean realizadas las pruebas del sistema.

Capítulo siete: es la presentación del marco teórico donde se muestran todos los conceptos necesarios para poder comprender la solución del desarrollo de prototipo de análisis y clasificación de información. En esta documentación se incluye la información del sistema de videovigilancia, así como el estudio del aprendizaje profundo y sus aplicaciones en redes neuronales.

En los siguientes capítulos se define el método que se utilizará para realizar la investigación, cuáles son los pasos para lograr implementar un prototipo de análisis y clasificación de información visual, la técnicas que se deben utilizar para lograr la recolección de información, posteriormente la manera en que se realizará el análisis de esta información, también se incluye un estudio

donde se expone la factibilidad del estudio y se define lo que se requiere para lograrlo de manera operativa, técnica y económica, en esta sección se detalla la información del porqué es factible realizar esta solución.

## 2. ANTECEDENTES

El monitoreo de cámaras de video es normalmente realizado por operadores humanos. La tarea de monitoreo según afirma Jimenez, (2010) “exige atención en tiempo real para detectar eventos y esto da como resultado baja precisión en su detención debido al cansancio, falta de observación y concentración” (p. 3), a veces por la cantidad de información o falta de concentración, se cometen errores, dando como resultado los sistemas no efectivos; por lo que han surgido investigaciones que pretenden sustituir este monitoreo o realizarlo de manera diferente para mejorar su eficiencia.

Tiwari y Verma (2015) presentan un artículo en donde proponen un método de detección de armas en imágenes utilizando segmentación basada en color y detección de puntos de interés Harris, de manera que puedan eliminar objetos que no son de interés. Para que cada uno de los puntos de interés detectados puedan dar la descripción del objeto donde si el 50 % del objeto encaja con las características de un arma pueda ser clasificada como tal. El método utilizado define que, sin importar la escala, rotación o variación pueda detectar una o varias imágenes dentro de una escena visual.

Relacionado con la detección de imágenes en videos se realizó una tesis llevada a cabo por el Tecnológico Nacional de México, Instituto Tecnológico de Celaya donde se ubicó una detección eficiente en entornos cerrados de cuchillos. Salgado, (2016) implemento la detección de imágenes utilizando contornos para identificar las armas blancas, donde el objetivo es comparar las imágenes para determinar si contienen las mismas características, realizando un entrenamiento

de inteligencia artificial con las 100 imágenes de cuchillos para lograr el reconocimiento.

Osorio, Olivera, Peña, Lopez Juárez, y Lefranc, (2016) realizaron un estudio donde se presenta un algoritmo de segmentación para detectar objetos en movimiento implementado en un sistema de vigilancia para realizar la supervisión de un estacionamiento. El cual consiste en la detección de objetos en movimiento mediante una secuencia de imágenes, donde se realiza la segmentación por modelado de fondo, que permite saber qué objetos no identificados se mueven, detectando el movimiento para luego identificar un objeto y se realiza una supervisión de este.

Li y Xie, (2019) presentan el artículo *A New Intelligent Video Surveillance Architecture*, en donde se muestra la implementación de una nueva arquitectura para el reconocimiento de imágenes mediante el uso de cámaras de vigilancia, la arquitectura es específicamente para el reconocimiento facial; su objetivo es detectar personas fugitivas y enviar alertas en el momento preciso. La arquitectura examina el video donde se obtienen las imágenes necesarias para ejecutar el reconocimiento de imágenes mediante interpolación, eliminación de ruido y otros métodos para finalmente codificar la imagen y compararla con las que se tienen codificadas en la nube. Esta arquitectura fue desarrollada en China por la Universidad Nacional de Defensa de la Tecnología y para realizar el reconocimiento facial se utilizó el algoritmo Viola-Jones y lograr resultados de alta precisión y detección en tiempo real.

En 2019 se realiza un estudio de reconocimiento de objetos mediante visión artificial utilizando openCV, Python, Numpy, redes neuronales convolucionales e histogramas. Para implementar este sistema se utiliza la operación de convolución realizada a una imagen digital. Se realizó el

preprocesamiento de imagen convirtiendo a escala de grises con el objetivo de reducir los costos computacionales y luego implementar métodos necesarios para la convolución de imágenes. Este sistema contiene una interfaz que permite ingresar cualquier imagen y aplicar diferentes filtros. (González, Zúñiga, Naola, Tito y Dongo, 2019).

En los últimos años el reconocimiento de imágenes es usado para la detección de objetos y personas debido a que su análisis mejora la precisión obtenida al realizar una revisión manual debido a que estos procesos son deficientes en comparación de los obtenidos por automatización, por esto sustituir este proceso por sistemas expertos en información es algo que está siendo utilizado. Estas soluciones han sido aplicadas a sistemas de videovigilancia, planteando arquitecturas, métodos de reconocimiento y manejo de imágenes para optimizar los recursos.



### **3. PLANTEAMIENTO DEL PROBLEMA**

La población de la ciudad de Guatemala diariamente presencia actos de peligro que suceden dentro de centros comerciales, comercios, restaurantes, y, cuando ocurre este tipo de hechos de manera frecuente, se cataloga esa zona como peligrosa. Según el análisis de seguridad ciudadana de Guatemala realizado por USAID con datos obtenidos de la Policía Nacional Civil de Guatemala en el año 2020 se determinó que el 70 % de los robos reportados al año estaban en los departamentos de Guatemala y Escuintla. Estos actos se ejecutaron el 38.3 % en viviendas, 54.4 % en comercios, 4.1 % en transporte público y un 1.8 % en iglesias. Debido a estas circunstancias las personas optan por contratar servicios de seguridad, cámaras y otros métodos para prevenir este tipo de acciones, incluso tomando estas medidas no garantizan que la vigilancia sea la adecuada.

La seguridad es vulnerable, aun cuando se posee el equipo necesario para realizar un control debido a que no se administran de manera correcta los recursos visuales, cuando se utilizan cámaras para cubrir un área, deben ser rigurosamente monitoreadas y ubicar la seguridad en el lugar correcto. Los sistemas tradicionales de videovigilancia que son monitoreados por personas son ineficientes por que la cantidad de cámaras en el sistema es mayor a la capacidad de una persona; cumplir esta tarea exige la atención completa de una persona para poder detectar eventos en tiempo real, esto da como resultado una tarea costosa en tiempo y errores humanos por falta de concentración.

Los recursos visuales proveen imágenes transmitidas en tiempo real que se almacenan de acuerdo a las políticas de la empresa, los registros se guardan

sin considerar su contenido, el manejo de la información impacta en tiempo al momento de utilizarlos e incluso de acuerdo a las políticas pueden ser eliminados, los datos almacenados no son útiles, porque la información se reduce a imágenes que deben ser interpretadas por algún tercero, los sistemas de seguridad visual son deficientes si para interpretar las escenas de peligro necesitan ser revisados nuevamente. El crecimiento de los contenidos de video en línea y fuera de línea exige ahora la indexación automática de contenido multimedia para facilitar la búsqueda y recuperación.

El análisis de la información en el momento adecuado podría mejorar el proceso de vigilancia y almacenamiento de evidencias, cuando se omite este proceso no se obtiene una conclusión, por ello se retrasan procesos legales o se concluyen al no poseer evidencia necesaria. Si el análisis de la información se ejecutara de manera eficiente, podría detectar actos inusuales antes de convertirse en un delito.

Esto lleva a plantear la pregunta principal de este estudio: ¿Cómo realizar la automatización del monitoreo de imágenes obtenidas de cámaras IP conectadas a una red local aplicando análisis y clasificación de información visual?

Para responder a esta interrogante se deberán contestar las siguientes preguntas auxiliares:

- ¿Permiten las redes neuronales la detección de actos inusuales en las imágenes obtenidas de cámaras IP conectadas a una red local?

- ¿Qué algoritmo de redes neuronales de detección de imágenes es eficiente para poder detectar actos inusuales en imágenes obtenidas de cámaras IP conectadas a una red local?
- ¿Qué variables permiten el reconocimiento correcto de imágenes para clasificar un acto inusual en el monitoreo de imágenes obtenidas de cámaras IP conectadas a una red local?



## 4. JUSTIFICACIÓN

Este trabajo corresponde a la línea de investigación dispositivos y sistemas para el desarrollo de sistemas que apliquen inteligencia artificial y *machine learning*. Esto a través de un sistema que permita la automatización del monitoreo de las imágenes extraídas de un circuito cerrado de televisión para realizar su análisis y clasificación de la información mediante el uso de técnicas de reconocimiento imágenes.

Los sistemas de seguridad deben ser monitoreados de la forma adecuada para impedir que ocurra cualquier acto delictivo. Con frecuencia en comercios ocurren estas situaciones; se observa al ingresar a diferentes lugares que existen imágenes de personas que han cometido delitos ahí. Esas imágenes fueron captadas por una cámara de seguridad, no fueron útiles debido a que no se evitó que sucediera o que vuelva a suceder.

La cantidad de información que almacenan los comercios por medio de las cámaras de vigilancia aumenta en cada segundo, la forma en que se guarda esta abarca espacio y se requiere demasiado tiempo para invertir en un monitoreo y posteriormente en una revisión manual. Los delitos no se encuentran hasta que se detectan faltantes en mercadería, una denuncia o algún otro hecho que afecte a una persona.

La implementación de un sistema que pueda detectar y clasificar la información de cuando se obtienen las imágenes de estos actos delictivos puede mejorar el manejo de la seguridad de algún comercio, enviando la seguridad al

lugar que lo necesita, para responder de manera eficiente y disminuir la falta de control en el establecimiento.

El reconocimiento de imágenes a través del uso de inteligencia artificial es una técnica que se está utilizando en diferentes áreas, los métodos de aprendizaje son útiles para que el sistema pueda tener su propio análisis y enfocarse en lo correcto para brindar un estudio de los datos que accedan de forma continua al sistema; para que el análisis y visualización sean rápidos y detectar situaciones de riesgo a tiempo y moderarlas.

Los sistemas de vigilancia son efectivos sin tener que utilizar operadores mediante los métodos de reconocimiento que se encargan de automatizar el monitoreo y alarmas, así como también notificar al usuario final. Esto se convierte efectivo debido a que el uso del internet permite el acceso a la información de los sistemas desde cualquier lugar, por ello revisar el estado del sistema no requiere que exista personal de monitoreo presente dentro de las instalaciones.

Al utilizar este sistema automatizado de videovigilancia los usuarios reducen los costos en personal encargado del monitoreo de las cámaras de seguridad y se benefician con alertas para mejorar el manejo de su seguridad y evitar así pérdidas materiales. Los administradores pueden hacer uso de la plataforma y visualizar los lugares vulnerables, por lo que realizarán una mejor administración de su seguridad en esos lugares.

## **5. OBJETIVOS**

### **5.1. General**

Implementar un prototipo que permita la automatización del monitoreo de imágenes obtenidas de cámaras IP conectadas a una red local aplicando el análisis y clasificación de información.

### **5.2. Específicos**

- Demostrar que el uso de redes neuronales mejora la detección de actos inusuales en la información visual obtenida de cámaras IP conectadas a una red local.
- Seleccionar el algoritmo de redes neuronales eficiente para detectar actos inusuales en el monitoreo de imágenes obtenidas de cámaras IP conectadas a una red local.
- Identificar las variables que permiten el reconocimiento correcto de imágenes obtenidas de cámaras IP conectadas a una red local para clasificar actos inusuales.



## **6. NECESIDADES A CUBRIR Y ESQUEMA DE LA SOLUCIÓN**

### **6.1. Necesidades a cubrir**

A partir de la implementación del prototipo que permita el análisis y clasificación de la información visual para la automatización del monitoreo de cámaras IP conectadas a una red local, se aplicará el uso de la inteligencia artificial para mejorar los procesos actuales con sistemas de videovigilancia. A pesar de que existen investigaciones relacionadas al uso de tecnología de reconocimiento facial y de imágenes, no existe un sistema en Guatemala que realice el monitoreo de cámaras de vigilancia por medio de la conexión a una red en donde no sea importante definir una marca específica.

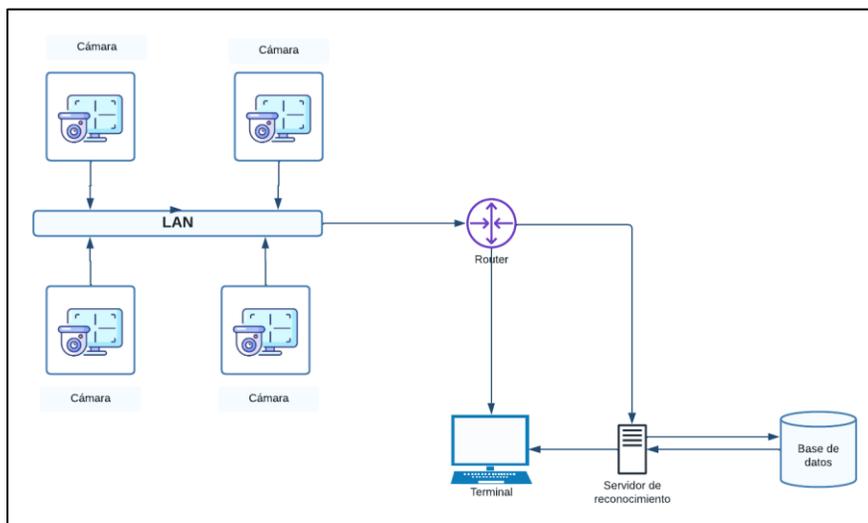
La necesidad de un proceso de automatización en la videovigilancia conectada a una red local es debido a que existen sistemas que pueden ser utilizados para este monitoreo, sin embargo, en la ciudad de Guatemala, este proceso se realiza de forma operativa en la mayor cantidad de comercios que no tienen acceso a un sistema que permita su automatización debido a los costos de comprar un sistema de videovigilancia. La propuesta de conectarse a un área local es que se pueda utilizar en diferentes dispositivos que permita obtener acceso a la información visual.

### **6.2. Esquema de solución**

Para la implementación del prototipo se utilizará la arquitectura que se muestra en la figura 1, donde se utilizarán cámaras IP las cuales necesitan ser configuradas para tener acceso a una red local y obtener las imágenes; la entrada

al sistema serán las imágenes obtenidas del sistema de videovigilancia, la información se traslada por medio de la red al servidor de reconocimiento de imágenes donde se realizarán las comparaciones mediante inteligencia artificial con el método de redes neuronales recurrentes. Esta información se guarda en una base de datos que interactúa de forma bidireccional con el servidor de reconocimiento de imágenes que es el encargado de proporcionar la salida que será enviada a un dispositivo conectado con el sistema general y el sistema de alarmas para informar al usuario final.

Figura 1. **Arquitectura general del sistema**



Fuente: elaboración propia, realizado con lucidchart.com.

A continuación, se describen los componentes clave de la arquitectura ilustrados en la figura 1:

- Cámaras IP: serán utilizadas para obtener las imágenes del sistema de videovigilancia.

- *Router*: para integrar el sistema de información necesitamos que las cámaras y el dispositivo que se utilizará estén conectados mediante la misma red de manera que puedan asegurar la conectividad y así trasladar la información a los demás dispositivos.
- LAN: las cámaras y dispositivos que manejan la información deben de estar dentro una red de área local.
- Base de datos: será utilizada para guardar la información de los pixeles de las imágenes para crear las comparaciones necesarias con la información almacenada de las imágenes con situaciones inusuales.
- Microservicios: se utilizarán dos microservicios para generar la salida de información, los cuales serán el servidor de codificación que obtiene las imágenes para procesarlas a pixeles y, el de reconocimiento de imágenes que permitirá realizar la comparación entre lo guardado en la base de datos y los nuevos datos.
- Alarmas: esta será la salida del sistema y serán activadas si se detecta alguna escena que deba ser reportada.
- Servidor de reconocimiento de imágenes: su objetivo es proporcionar las salidas utilizando la información de la base de datos y el entrenamiento realizado con redes neuronales recurrentes.



## **7. MARCO TEÓRICO**

### **7.1. Sistema de videovigilancia**

A continuación, se hace una breve descripción acerca de los sistemas de videovigilancia.

#### **7.1.1. Definición**

Un sistema de videovigilancia, según Mata, (2010) es “la supervisión, el control y el eventual registro de actividad física dentro de un local, espacio o ambiente en general” (p. 11). La función de un sistema de vigilancia es proveer los recursos necesarios para poder efectuar un control en base a lo que se logra en la observación de los hechos, por eso se define como un sistema de control y supervisión. También se hace énfasis a lo que es el registro de actividad y con en base a esto tomamos en consideración que toda la información captada por un sistema de videovigilancia está siendo almacenada. Estos sistemas están delimitados con base en su capacidad de acuerdo con la cantidad de cámaras que se puedan tener en el espacio en que se realiza el monitoreo.

Los sistemas de videovigilancia están compuestos por una o varias cámaras que se encuentren conectadas al sistema, estos pueden transmitir las imágenes en monitores, televisiones en tiempo real y de acuerdo a las necesidades del usuario son almacenados; las cámaras que se utilizan son definidas por el usuario para determinar su posición y de qué manera deben ser colocadas. En estos sistemas de videovigilancia con regularidad se tienen salas de control donde pueden ser visualizadas las imágenes por un operador

encargado de realizar el monitoreo, donde tiene opción de cambiar en diferentes vistas así como obtener la configuración requerida para poder realizar su trabajo de supervisión de manera correcta.

### **7.1.2. Videovigilancia IP**

La videovigilancia ha tenido cambios. En los inicios se observaban sistemas realizados únicamente con sistemas análogos y, estos eran rígidos, su forma de utilizar era compleja y se tenían que utilizar diferentes dispositivos para poder obtener la información. Dado estos cambios surgió la videovigilancia utilizando cámara IP debido a que esta tecnología es más flexible y podemos obtener la información mediante servicios web o cualquier dispositivo que se encuentre conectado a una red que puede ser una red LAN o una red WAN.

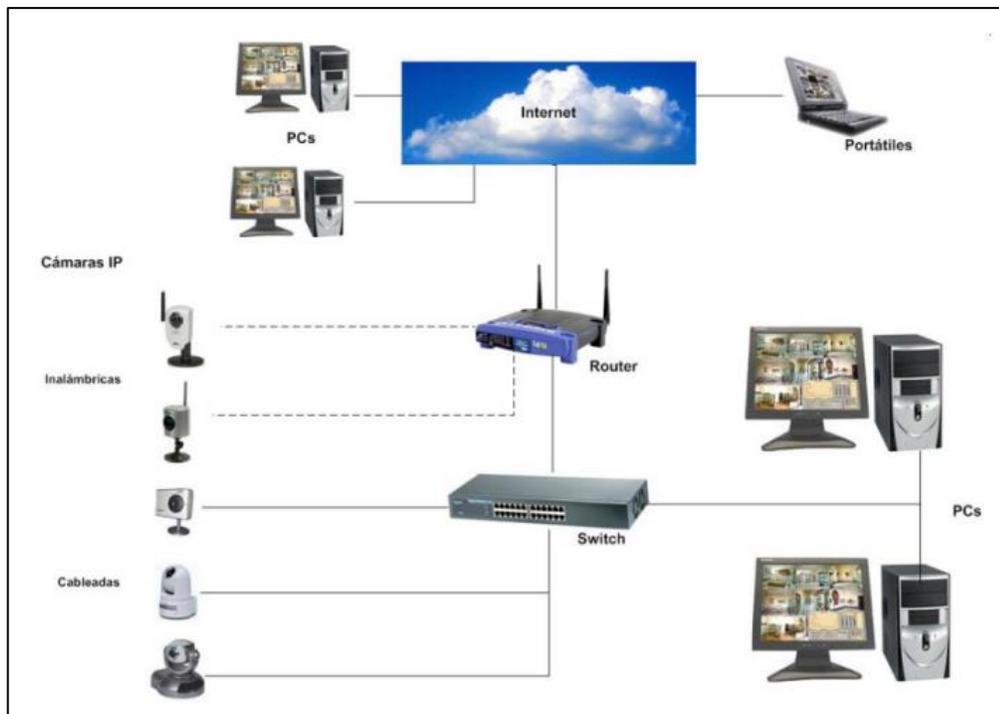
La forma de obtener la información visual es por medio de las conexiones realizadas mediante la red que pueden ser de forma cableada o inalámbrica. Este tipo de conexiones facilita que se pueda realizar un monitoreo en tiempo real, así como conectarse a otros dispositivos que permitan la transmisión, almacenamiento y gestión de las alertas que debe enviar el equipo encargado.

Los sistemas de videovigilancia IP han mejorado en varios sentidos en comparación con los análogos utilizados tradicionalmente, entre ellos se encuentra la facilidad con la accesibilidad por medio de la red, la calidad de los videos, la facilidad para integrar cámaras, la escalabilidad, flexibilidad y mejora en la rentabilidad de obtener un sistema de videovigilancia.

### 7.1.3. Arquitectura

La videovigilancia IP incluye cámaras IP que son las encargadas de captar y transmitir la información de audio y video a través de una red IP a otros dispositivos. Esto es realizado mediante el direccionamiento de IP, utilizando un servidor web y protocolos de *streaming* de video. Lo que permite al usuario final visualizar y gestionar su video de forma remota o local, facilitando que los usuarios puedan mantenerse enlazados al sistema mediante diferentes dispositivos que se encuentren conectados a una conexión de internet. Para definir la arquitectura de un sistema de videovigilancia IP se representa gráficamente en la figura 2.

Figura 2. Esquema básico de un sistema de Videovigilancia IP

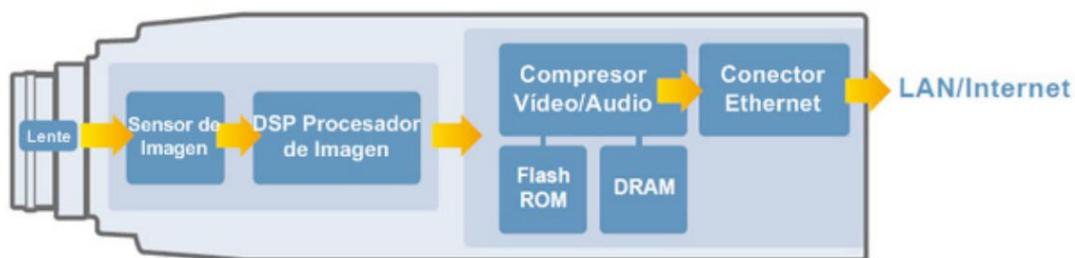


Fuente: Mata. (2010). *Vigilancia CCTV usando video IP*.

#### 7.1.4. Cámara IP

Podemos definir una cámara IP de acuerdo con Mata, (2010) como “un dispositivo que capta y transmite una señal de audio/vídeo digital través de una red IP estándar a otros dispositivos de red, tales como un PC o un teléfono 3G” (p. 17). Una cámara IP es un dispositivo capaz de poder obtener imágenes para luego procesarlas, comprimirlas y transmitir estos datos por medio de la dirección IP mediante su conexión a la red. La forma en que estas procesan las imágenes puede mostrarse en la figura 3, donde se visualiza las partes y procesos que realiza una cámara IP para poder obtener la información visual transmitida al sistema de videovigilancia.

Figura 3. Esquema básico de una cámara de red



Fuente: Mata. (2010). *Vigilancia CCTV usando video IP*.

Existe una variedad de este tipo de cámaras; podemos encontrar para interiores, exteriores, fijas o domo fijas hasta PTZ. Estas cámaras nos permiten poder monitorear lo que necesitamos de acuerdo con las necesidades y recursos disponibles.

### **7.1.5. Transmisión de datos**

En un sistema de videovigilancia es importante contar con la transmisión de datos, debido a que “Es la técnica que se emplea para transportar información que se genera, procesa y almacena en los sistemas de cómputo” (Pérez, 2010, p. 17). Un sistema de transmisión de datos es el encargado de transportar la información que se desea enviar, en donde podemos encontrar diferentes tipos de información en un sistema de videovigilancia, nos importa que las imágenes puedan ser procesadas y recibidas posteriormente por un servidor, sin embargo, esto no sería posible si no existieran redes avanzadas de computación que son el medio por el que se comparte y transfieren datos a diferentes áreas geográficas de manera rápida sin importar el volumen de estos.

Los sistemas de transmisión de datos pueden ser complejos o sencillos, utilizando diferentes redes para trasladar la información a un dispositivo periférico. Para completar este proceso se utilizan módems o redes complejas para obtener información de un centro de datos o simplemente de una red local que puede interconectar los diferentes dispositivos para una casa, empresa, institución, etc.

Estos sistemas deben abarcar ciertos requisitos debido a que la forma en que transmiten la información es por medio de señales de datos, por lo que si una señal es enviada de manera incorrecta puede cambiar el significado por completo. Para lograr la transmisión de datos con un control adecuado de flujo de datos, codificación de información y corrección de datos es necesario tener protocolos de comunicación los cuales aseguran que sea la correcta con un sistema de computación. Se consideran los protocolos para transferencia debido a que estos establecen la comunicación organizada para lograr que la

información fluya de manera apropiada y ordenada para que sea coherente y significativa en el intercambio de información.

#### **7.1.6. Monitoreo de sistemas de vigilancia**

Las imágenes captadas deben ser transmitidas hacia dispositivos que permitan reproducir y monitorear para lograr la supervisión, interpretación y un control por el personal destinado a esta responsabilidad, según Fernandez, (2013) “Los dispositivos de reproducción y monitorización son transductores de video que realizan la función de recuperar la señal eléctrica del video y volverla a convertir en imagen” (p. 166). Para realizar esta monitorización de una manera operativa se utilizan diferentes tipos de dispositivos de reproducción en los que podemos encontrar monitores análogos y monitores digitales que actualmente son los más utilizados.

Para realizar monitoreo de cámaras de vigilancia se puede definir un sistema de monitorización múltiple como el que tiene un monitor por cada cámara conectada al sistema de seguridad, cuando se utiliza este tipo de sistema de seguridad suele ser demasiado costoso en recursos y personal, se necesita un alto presupuesto para comprar la cantidad de monitores y es complicado para una persona poder observar varias pantallas al mismo tiempo. Como alternativa se tiene el sistema de monitorización de red, el cual consiste en que las imágenes captadas puedan ser transmitidas a un dispositivo que se encuentre conectado, en este permitirá observar la cantidad de cámaras utilizadas en un solo dispositivo o de acuerdo con la configuración requerida, este sistema facilita el monitoreo porque se puede realizar en diferentes lugares mientras se tenga la red necesaria.

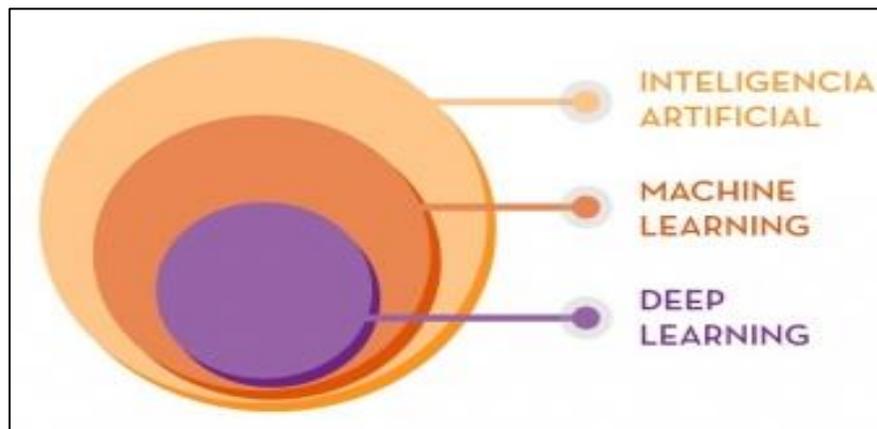
## 7.2. *Deep learning*

En el siguiente inciso se describe qué es el *deep learning*.

### 7.2.1. Definición

*Deep learning* se traduce como aprendizaje profundo que a su vez forma parte del aprendizaje automático y este es parte del campo de inteligencia artificial; esto se puede visualizar en el diagrama de la figura 4, donde se muestra que el *deep learning* es la capa que está en el interior de ambos campos, por lo que este encapsula los métodos utilizados en inteligencia artificial y *machine learning*.

Figura 4. **Relación entre inteligencia artificial, *machine learning* y *deep learning***



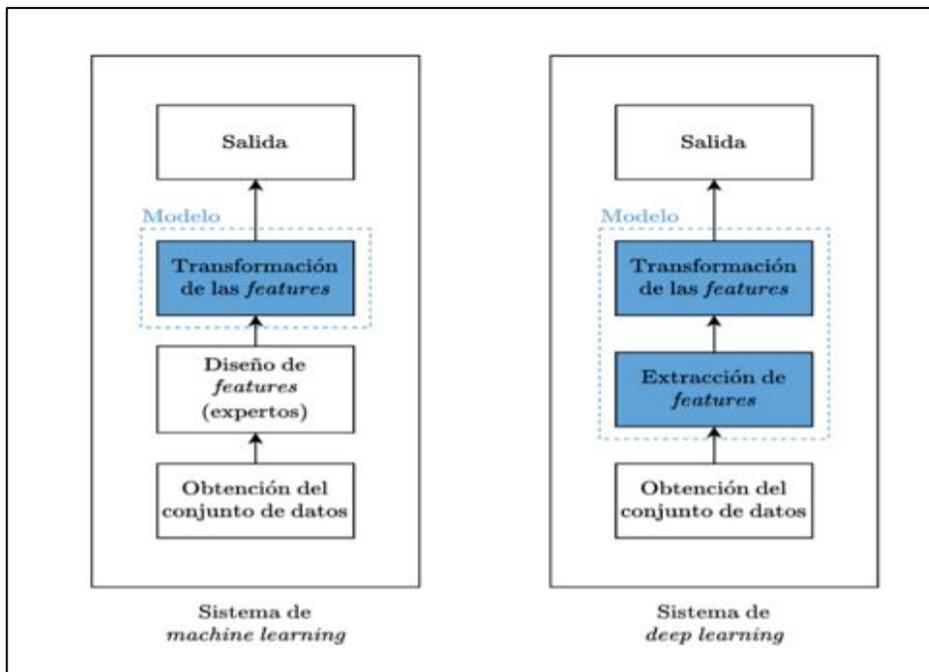
Fuente: Ministerio de asuntos económicos y transformación digital. (2020). *Representación gráfica de la relación entre inteligencia artificial, machine learning y deep learning*, p. 8.

Consultado el 05 de mayo de 2022. Recuperado de:

[https://datos.gob.es/sites/default/files/doc/file/tecnologias\\_emergentes\\_y\\_opendata\\_inteligencia\\_artificial-\\_final\\_0.pdf](https://datos.gob.es/sites/default/files/doc/file/tecnologias_emergentes_y_opendata_inteligencia_artificial-_final_0.pdf)

El *machine learning* consiste en un conjunto de datos que debe ser ajustado a un modelo para resolver un problema, los datos se obtienen de un estudio realizado por expertos que definen características del modelo permitiendo elegir el método adecuado y obtener soluciones. En este punto el *deep learning* no necesita que se realice el estudio previo para determinar los datos que se deben utilizar, por lo que este método toma los datos originales de entrada para poder resolver el problema, consiste en un aprendizaje de transformación para poder dar las salidas esperadas, en este punto podemos tomar en consideración que el *deep learning* define sus propias características con base en el aprendizaje. Estas diferencias las podemos observar en la figura 5, que muestra el proceso que realiza el *machine learning* y el *deep learning*.

Figura 5. **Diferencias entre *machine learning* y *deep learning***



Fuente: Pérez y Gegúndez. (2021). *Deep Learning*.

Para lograr el autoaprendizaje en el *deep learning* se necesitan varias capas dentro de una red neuronal para tener las combinaciones necesarias y poder encontrar la solución propuesta; de ahí es donde nace el término de “profundo”, mientras que en el *machine learning* no es necesario tener una red neuronal multicapa para definir la salida esperada debido a que es parte de la definición del sistema configurar las características necesarias para obtener los resultados. Para lograr el aprendizaje profundo se obtienen las características de los datos y en cada capa distingue la forma de encontrar características en cada nivel de mayor prioridad para obtener la salida, por ello el aprendizaje profundo está definido por dos etapas: una de extracción y la otra de transformación.

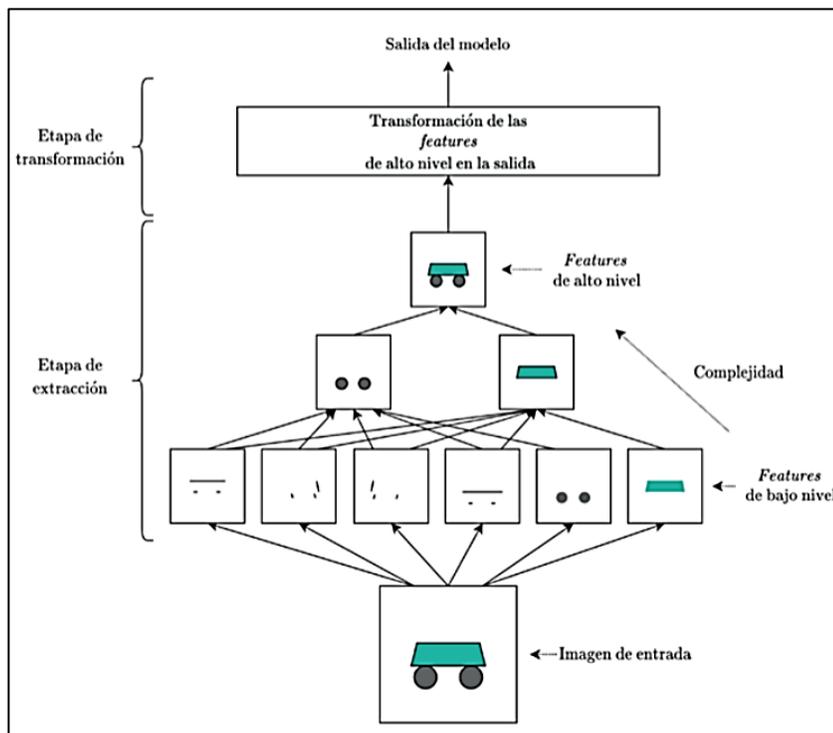
Las ventajas de utilizar *deep learning* en comparación con *machine learning* son que facilita el desarrollo de modelos y evita los costos extras que produce el análisis de características a definir para el funcionamiento correcto, aparte de ello poseen limitantes como la cantidad de información analizada para obtener las características que puede ser difícil de capturar, si la información está cambiando con el tiempo, mientras que el *deep learning* puede cambiar las características de acuerdo a los datos utilizados para dar los resultados esperados. A pesar de que utilizar *deep learning* ahorra fases del diseño para el análisis de información, podemos encontrar que estos métodos son más costosos en forma tecnológica, por ello exigen equipo con recursos demandantes en capacidad de procesamiento para manejarlos, lo que no suele ser tan utilizado por la necesidad de comprar un equipo costoso con los recursos requeridos.

### **7.2.2. Componentes**

El *deep learning* está compuesto por dos etapas, la de extracción y la de transformación. La primera debe extraer los datos tal como se reciben en el sistema obteniendo las características mientras que, la segunda, realizará

transformaciones para encontrar las características de más alto nivel que permiten una salida correcta. Este proceso lo podemos observar en la figura 6.

Figura 6. **Modelo *deep learning* con sus diferentes etapas**



Fuente: Pérez y Gegúndez. (2021). *Deep Learning*.

En esta figura se puede observar de qué manera se forman las capas que van analizando y clasificando las características de acuerdo con las que considere que tienen más valor y descarta las que no son necesarias. Para obtener el correcto procesamiento de la información, que en este esquema es una imagen, se puede observar de qué manera descarta las características que no tiene valor y así poder retornar la imagen esperada por el usuario final. Este esquema simula una red neuronal convolucional. Estas redes neuronales son las encargadas del procesamiento de imágenes, sin embargo, la forma en que

funciona el *deep learning*, utilizando estas dos fases, es parecido para las otras aplicaciones en que se pueda utilizar este aprendizaje.

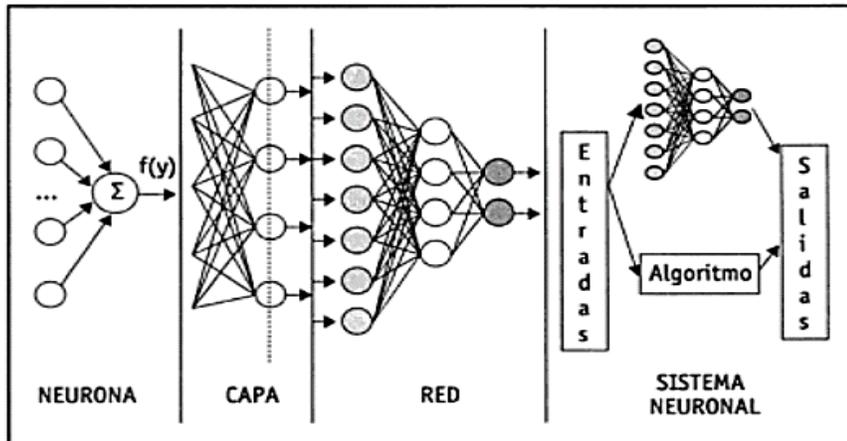
### **7.3. Redes neuronales**

La definición acerca de una red neuronal se describe a continuación.

#### **7.3.1. Definición**

Podemos definir a las redes neuronales de acuerdo con Flórez y Fernández, (2008) quienes indican que estas “Tratan de emular el comportamiento del cerebro humano, caracterizado por el aprendizaje a través de la experiencia y la extracción de conocimiento a partir de un conjunto de datos” (p. 11). Las redes neuronales intentan imitar cómo funcionan las redes neuronales biológicas, las cuales a partir de una entrada de datos y su análisis intentan darnos una salida. Para lograr esto es necesario que la red neuronal tenga una estructura con la que pueda simular el comportamiento para luego, mediante la modelación de procesos matemáticos, se pueda obtener la salida. Se puede visualizar cómo se define una red neuronal en la figura 7.

Figura 7. Estructura jerárquica de un sistema basado en RNA



Fuente: Flórez y Fernández. (2008). *Las redes neuronales artificiales, Fundamentos teóricos y aplicaciones prácticas.*

### 7.3.2. Tipos de aprendizaje

A continuación, en los siguientes incisos se describen los diferentes tipos de aprendizaje que existen.

#### 7.3.2.1. Aprendizaje supervisado

En el aprendizaje supervisado según Rouhiainen, (2018) se “Usan datos que ya han sido etiquetados u organizados previamente para indicar cómo tendría que ser categorizada la nueva información” (p.20), cuando se utiliza este tipo de aprendizaje es necesario que una persona pueda definir la manera correcta en que debe de comportarse el algoritmo, se define supervisado por que necesita que se le den los parámetros para establecer si lo que se está realizando es correcto, mediante las respuestas que obtiene definirá sus entradas correctas e ira aprendiendo de las retroalimentaciones. Este tipo de aprendizaje es

comúnmente utilizado para clasificación y regresión debido a que en estos casos se tiene definida una variable objetivo que permitirá al algoritmo definir si su salida es correcta.

### **7.3.2.2. Aprendizaje no supervisado**

En el aprendizaje no supervisado según Rouhiainen, (2018) “No usan ningún dato etiquetado u organizado previamente para indicar como tendría que ser categorizada la nueva información” (p.20), cuando se tiene un aprendizaje no supervisado no es necesario que alguien intervenga en el sistema porque su funcionamiento consiste en que, el sistema pueda aprender de manera independiente para clasificar la información. Actúa por medio de la información con la que se alimenta el algoritmo, incluso este tipo de sistemas suelen alimentarse de la misma información utilizada previamente para determinar las salidas del sistema.

### **7.3.3. Aplicaciones**

Las aplicaciones son las diferentes formas en las que se pueden utilizar las redes neuronales.

#### **7.3.3.1. Concepto de procesamiento de imágenes**

Los humanos hemos realizado el procesamiento de imágenes desde la visualización y analizamos su contenido, las personas han realizado análisis con base en la observación; sin embargo, una persona no es capaz de observar o tomar en consideración detalles como si lo hiciera una computadora, ambos sistemas son independientes entre sí, porque las redes neuronales se basan en poder replicar el comportamiento que tienen las neuronas biológicas. Vilet,

(2005) afirma “El campo del procesamiento digital de imágenes está construido sobre bases matemáticas y probabilísticas, pero la intuición y análisis humanos juegan un papel importante” (p.15). La forma en que ha evolucionado la tecnología y las ciencias ha logrado que se puedan realizar este tipo de sistemas, con base en diferentes algoritmos de inteligencia artificial mediante modelos matemáticos.

#### **7.3.4. Redes neuronales convolucionales**

En el siguiente inciso se da una breve definición acerca de las redes neuronales convencionales.

##### **7.3.4.1. Definición**

Las redes neuronales convolucionales contienen la misma estructura que las demás redes, pero estas destacan con su particularidad que consiste en que normalmente son utilizadas para la clasificación de imágenes y, según afirma Torres (2018), “Un rasgo diferencial de las CNN es que hacen la suposición explícita de que las entradas son imágenes, cosa que nos permite codificar ciertas propiedades en la arquitectura para reconocer elementos concretos en las imágenes” (p. 147). Por lo que este tipo de redes neuronales tomará en consideración los patrones que pueden identificar una imagen de acuerdo con lo que estamos buscando clasificar; como lo podrían ser líneas, bordes, texturas o formas similares que se intentan identificar en una imagen, a través de estos rasgos se definen las capas que debe tener la red neuronal.

Las redes neuronales convolucionales realizan la descomposición de los elementos que recibe para analizar estas características. Existen diferentes

técnicas que permiten por medio de filtros y otras configuraciones establecer bordes para obtener mejor la representación de las imágenes.

#### **7.3.4.2. Detección de bordes en imágenes**

En el tratamiento de imágenes uno de los principales métodos que se debe revisar es la detección de bordes en imágenes debido a que estos identifican los cambios que se desarrollan de manera inesperada en una imagen, los cuales muestran información acerca de la imagen y su relación con los patrones establecidos para el reconocimiento. En los bordes se encuentra la información relevante de una imagen, debido a que allí están las características que pueden representar algo importante en las comparaciones de la red neuronal.

Para la extracción de bordes en una imagen se aplican diferentes tipos de filtros que permiten encontrar las regiones con bordes. Estos pueden clasificarse de acuerdo con la anchura, ángulo y coordenadas. Para lograr observar estos bordes se realiza mediante la aplicación de filtros y funciones matemáticas que mediante cada ciclo de análisis aclarará qué tan representativo es este borde.

#### **7.3.5. Redes neuronales recurrentes**

A continuación, se da una breve definición acerca de las redes neuronales recurrentes.

##### **7.3.5.1. Definición**

Las redes neuronales recurrentes según Torres (2018), “Son una clase de redes para analizar datos de series temporales permitiendo tratar la dimensión de tiempo” (p. 58). Estas redes actúan de manera que siempre incluyen

conexiones que apuntan hacia atrás de manera que se basan en salidas anteriores, su resultado se basará de acuerdo con lo que ha pasado, lo importante en este tipo de redes neuronales es definir lo que sucede en un instante de tiempo, poder explicar lo que sucedió en un periodo de tiempo.

#### **7.3.5.2. Memoria a corto plazo (LSTM)**

Al considerar las redes neuronales recurrentes se consideran los diferentes algoritmos definidos para realizar el análisis de datos, entre estos algoritmos definimos la memoria a corto plazo y según Torres, (2018) “básicamente amplían su memoria para aprender de experiencias importantes que han pasado hace mucho tiempo” (p. 60). Estas redes neuronales recurrentes pueden analizar un periodo de tiempo debido a que siempre estarán utilizando la información que tiene guardada en su memoria de iteraciones anteriores. Este tipo de redes neuronales puede realizar una decisión para determinar si almacena o elimina la información, esta decisión la tomará con base en la información que está recibiendo, este algoritmo aprende a partir de si la información guardada es útil en los casos detectados futuramente.

## **8. PROPUESTA DE ÍNDICE DE CONTENIDO**

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES

LISTA DE SÍMBOLOS

GLOSARIO

RESUMEN

PLANTEAMIENTO DEL PROBLEMA

OBJETIVOS

RESUMEN DE MARCO METODOLÓGICO

INTRODUCCIÓN

1. ANTECEDENTES

2. JUSTIFICACIÓN

3. ALCANCES

3.1. Perspectiva investigativa

3.2. Perspectiva técnica

3.3. Perspectiva de resultados

4. MARCO TEÓRICO

4.1. Sistema de videovigilancia

4.1.1. Definición

4.1.2. Videovigilancia IP

4.1.3. Arquitectura

4.1.4. Cámara IP

- 4.1.5. Transmisión de datos
- 4.1.6. Monitoreo de sistemas de vigilancia
- 4.2. *Deep learning*
  - 4.2.1. Definición
  - 4.2.2. Componentes
- 4.3. Redes neuronales
  - 4.3.1. Definición
  - 4.3.2. Tipos de aprendizaje
  - 4.3.3. Aplicaciones
  - 4.3.4. Redes neuronales convolucionales
  - 4.3.5. Redes neuronales recurrentes

## 5. PRESENTACION DE RESULTADOS

## 6. DISCUSIÓN DE RESULTADOS

CONCLUSIONES

RECOMENDACIONES

REFERENCIAS

ANEXOS

## **9. METODOLOGÍA**

Se realizará una investigación cuantitativa. La solución propuesta implementará una automatización en la forma en que se realiza actualmente el monitoreo de cámaras de vigilancia, logrando que se tenga una detección rápida y los usuarios no tengan que pasar largas horas frente al sistema de vigilancia para determinar si ha ocurrido algo inusual y avisar mediante alertas. El objetivo es poder demostrar que los sistemas de inteligencia artificial reducen costos operativos y proveen exactitud.

Esta investigación tecnológica demostrará que la creación y aplicación de modelos de inteligencia artificial pueden mejorar los sistemas que requieren de recurso humano y pueden aumentar su productividad.

### **9.1. Diseño de investigación**

Es experimental debido a que se realizará una revisión de cómo se encuentran actualmente los sistemas de vigilancia y la eficiencia que tienen estos para, posteriormente realizar un estudio donde se pueda realizar una comparación de la detección de los actos inusuales que pueden ser detectados mediante las redes neuronales.

### **9.2. Variables de estudio**

A continuación, en la tabla I se describen las variables del estudio.

Tabla I. **Variables de estudio**

<b>Variable</b>	<b>Definición</b>	<b>Subvariable</b>	<b>Indicadores</b>
Eficiencia de algoritmos de redes neuronales en reconocimiento de imágenes	Capacidad del algoritmo para detectar actos inusuales.	<ul style="list-style-type: none"> <li>• Curva de aprendizaje</li> <li>• Recursos tecnológicos</li> <li>• Imágenes</li> </ul>	<ul style="list-style-type: none"> <li>• Dificultad del aprendizaje</li> <li>• Cantidad de aciertos por el algoritmo</li> <li>• Facilidad de implementación</li> </ul>
Eficiencia en detección de acto inusual en sistema de monitoreo	Capacidad del sistema de monitoreo para detectar y alertar cuando ocurre un acto inusual.	<ul style="list-style-type: none"> <li>• Algoritmos de reconocimiento de imágenes</li> <li>• Operadores</li> </ul>	<ul style="list-style-type: none"> <li>• Tiempo en monitoreo</li> <li>• Tiempo que tardan las alertas</li> </ul>
Condiciones de automatización de monitoreo	Efectos de utilizar sistemas de reconocimiento de imágenes para reemplazar el monitoreo operativo.	<ul style="list-style-type: none"> <li>• Ventajas de reconocimiento</li> <li>• Desventajas de reconocimiento</li> </ul>	<ul style="list-style-type: none"> <li>• Costo</li> <li>• Adaptabilidad</li> <li>• Facilidad de implementación</li> <li>• Recursos necesarios</li> </ul>

Fuente: elaboración propia, realizado con Word.

### 9.3. Alcances

- Perspectiva investigativa

Esta investigación tiene la perspectiva de investigación exploratoria descriptiva.

Los alcances investigativos surgen del desarrollo de un prototipo que permita la clasificación de información integrado a un sistema de video vigilancia. El sistema debe recolectar la información desde cámaras de videovigilancia hasta poder detectar actos inusuales y dar aviso al usuario que utiliza el sistema.

El alcance descriptivo es:

- La investigación define los algoritmos que se deben utilizar para determinar la efectividad de las redes neuronales en la detección de imágenes.

El alcance exploratorio define:

- La investigación deberá considerar los recursos, procedimientos y componentes para poder definir un sistema que obtenga las imágenes de cámaras IP de un sistema de videovigilancia, algoritmos de reconocimiento de imágenes.

- **Perspectiva técnica**

Un sistema de clasificación de información visual es aquel que obtiene imágenes y puede detectar patrones definidos por el usuario, con el fin de disminuir el uso de recursos operativos y tecnológicos.

Los aspectos técnicos que cubrirá son:

- Clasificación de imágenes utilizando reconocimiento de imágenes mediante redes neuronales recurrentes.
- Elección de variables que permitan el correcto funcionamiento del reconocimiento de imágenes.

- Estudio económico de los costos de los dispositivos requeridos para que un individuo con bajo presupuesto implemente su sistema de videovigilancia en base al prototipo.
- Envío de alertas al usuario final.
- Perspectiva de resultados

La clasificación de información visual está enfocada en permitir el monitoreo mediante la clasificación y optimización de imágenes dentro de un sistema de videovigilancia que debe estar conectado a una red local.

Los resultados son:

- Ejecución de escenarios de prueba donde se clasifique imágenes, detectando coincidencias en las comparaciones con redes neuronales recurrentes.
- Recolección de resultados de ambos métodos para determinar la eficiencia del monitoreo utilizando inteligencia artificial contra la eficiencia obtenida utilizando métodos tradicionales.
- Demostración de la efectividad de las variables seleccionadas para el reconocimiento de imágenes.
- Demostración de la efectividad de automatizar el monitoreo de cámaras de seguridad.

- Demostración de los recursos y costos que se requieren para instalar un sistema de seguridad en una vivienda utilizando el prototipo.
- Evaluación de la factibilidad de un individuo con bajo presupuesto para para instalar un sistema de seguridad en una vivienda utilizando el prototipo.

#### **9.4. Procedimiento metodológico**

Se definen cada una de las fases involucradas para la elaboración de la investigación desde el inicio hasta la ejecución y pruebas finalizadas.

##### **9.4.1. Observación documental**

Se realizará una observación técnica mediante la realización de un estudio en que se pueda analizar una población que cuente con los requisitos para ser un agente de seguridad, con el fin de que pueda visualizar videos y posteriormente pueda detectar los actos inusuales vistos; se tratará de obtener detalles acerca de lo visualizado, así también se permitirá realizar revisiones en el video y cronometrar el tiempo utilizado para detectar esos detalles y encontrar los momentos con actos inusuales.

##### **9.4.2. Investigación documental**

Se realizará una investigación para obtener información acerca de:

- La manera en que funcionan los algoritmos de reconocimiento de imágenes actuales y de qué manera pueden ser aplicados al reconocimiento de actos inusuales en el tiempo.

- La manera adecuada de obtener las imágenes que se utilizaran para alimentar la base de datos del algoritmo.
- La manera en que se evaluará la efectividad de los algoritmos a probar.

#### **9.4.3. Decisión con base a conocimientos**

Utilizando los conocimientos adquiridos en la fase II se planea empezar a definir los siguientes componentes que servirán para el desarrollo del prototipo:

- Definición, posición y ángulo en que se colocarán las cámaras que se utilizarán para el monitoreo.
- Selección de librerías a utilizar para obtener las imágenes y entrenar la red neuronal.
- Selección de métodos a utilizar para el etiquetado de imágenes.
- Selección de algoritmos a probar para el reconocimiento de imágenes.

#### **9.4.4. Implementación de algoritmos de reconocimiento**

Con base en lo analizado durante la fase anterior se iniciará con la conexión de las cámaras al sistema partiendo con el desarrollo del servicio de reconocimiento de imágenes con los diferentes algoritmos existentes; donde se definirá el algoritmo más eficiente que será el utilizado en el prototipo.

#### **9.4.5. Implementación de prototipo**

Se iniciará con el desarrollo del prototipo que debe contar con los siguientes requerimientos:

- Configuración de la conexión de cámara IP al sistema de reconocimiento.
- Sistema de reconocimiento de imágenes utilizando el método más eficiente encontrado en la fase anterior.
- Alertas que se enviarán al usuario final con base al reconocimiento.

Se realizarán pruebas con el método de reconocimiento para obtener las alertas necesarias con los actos inusuales detectados.

#### **9.4.6. Análisis de resultados**

Con base en los datos recolectados de métodos convencionales y las pruebas ejecutadas en la implementación del prototipo se deberán de definir las siguientes variables:

- Porcentaje de eficiencia
- Tiempo para detectar actos inusuales
- Tiempo para obtener información
- Cantidad de actos inusuales detectados
- Costo de implementar el prototipo para un individuo en su vivienda

## 9.5. Técnicas de recolección de información

Para realizar la automatización del monitoreo de cámaras IP en una vivienda se recolectará información por medio de las siguientes técnicas:

- Revisión documental y bibliográfica: para estudiar los algoritmos que actualmente cumplen la función de reconocimiento de imágenes, la manera en que pueden operar para implementar la automatización del monitoreo. Para ello se realizará una revisión de libros, artículos científicos, tesis y cursos obteniendo la información necesaria para realizar la implementación de redes neuronales y poder determinar la eficiencia de los sistemas actuales e implementarlo en el sistema desarrollado.
- Observación documental: se realizarán pruebas con personas que cumplan con los requisitos de ser personal de monitoreo para determinar la eficiencia actual de los sistemas monitoreados por una persona. También se realizarán preguntas acerca de los aspectos del monitoreo operativo.

Observación experimental: se realizará una recolección de datos de acuerdo con lo que se observe en las pruebas controladas realizadas con cada uno de los algoritmos de reconocimiento de imágenes donde se recolectará información de tiempos, recursos y eficiencia.

## 10. TÉCNICA DE ANÁLISIS DE INFORMACIÓN

Se realizará el análisis de la información de lo recolectado con las técnicas descritas para lograr los objetivos de la investigación. Se utilizarán las siguientes técnicas de análisis de información:

- **Etiquetado de imágenes:** se realizará la clasificación de las imágenes que contengan actos inusuales que serán detectados por el algoritmo de inteligencia artificial para determinar la información que es útil para el entrenamiento de las redes neuronales, por medio de este proceso se analizará la información que deben utilizar los algoritmos.
- **Análisis del módulo de videovigilancia:** se revisará la información obtenida de las entrevistas para clasificar las etapas de las que depende de un monitoreo y realizar la estructura del sistema, visualizar los puntos débiles y fuertes del monitoreo operativo.
- **Redes neuronales:** se aplicarán los algoritmos de redes neuronales en un sistema de monitoreo de cámara de vigilancia para validar la información recolectada de los algoritmos y validar sus eficiencias determinadas por valores de precisión obtenidos de los mismos.
- **Métodos estadísticos:** se utilizará la matriz de confusión para obtener la eficiencia de los algoritmos aplicados. Mediante el estudio estadístico se podrá determinar el margen de error en cada uno de los algoritmos, así como definir la eficiencia contra el método operativo.

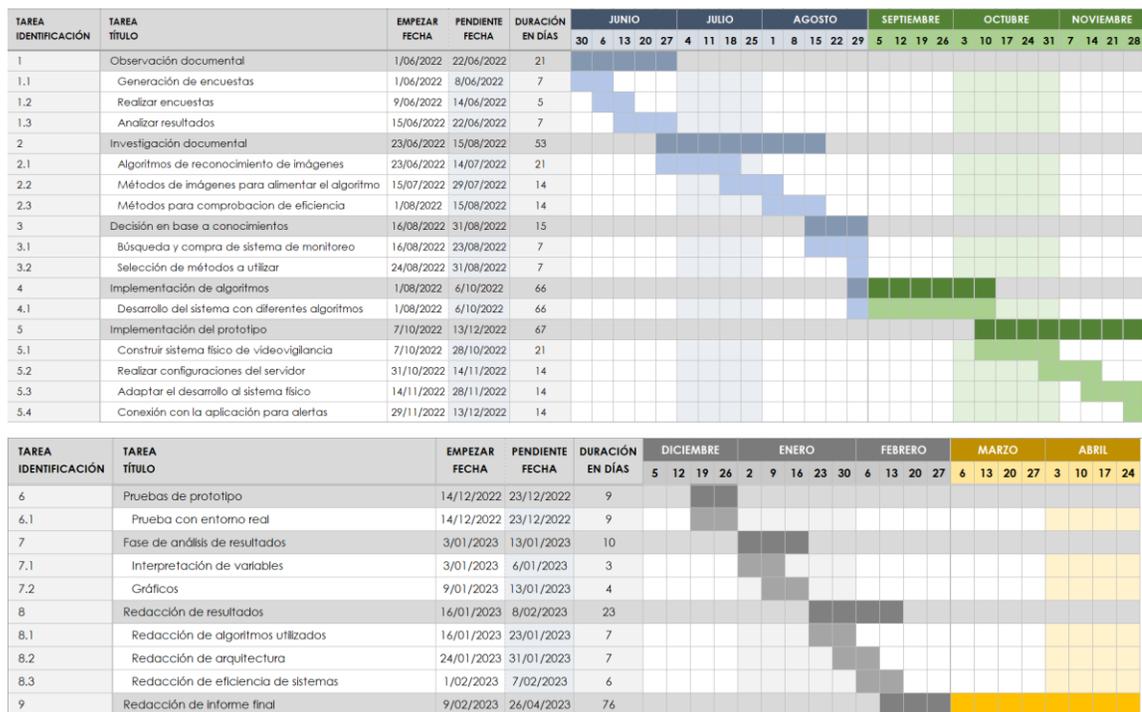
- Gráficos de comparación: se realizarán comparaciones entre las variables definidas para cada uno de los métodos implementados, así como la información obtenida de cada uno de los algoritmos aplicados para visualizar el rendimiento de los algoritmos de inteligencia artificial y método operativo.

Programación y optimización matemática: se realizará un estudio de costo beneficio para determinar si es factible económicamente para un individuo instalar el prototipo y, si logra reducir los costos con base a otros sistemas existentes.

## 11. CRONOGRAMA

Para el desarrollo de la investigación se presenta el orden en que se aplicará la metodología y el tiempo que tardará cada una de las fases de manera cronológica:

Figura 8. Cronograma



Fuente: elaboración propia, realizado con Excel.



## **12. FACTIBILIDAD DEL ESTUDIO**

Para realizar el estudio es necesario evaluar los recursos necesarios de su ejecución considerando los diferentes campos que son el recurso operativo, técnico y económico.

### **12.1. Factibilidad operativa**

Para el desarrollo del prototipo de análisis y clasificación de información visual se requiere de los siguientes componentes:

- Disponibilidad de personas para realizar el estudio de eficiencia de métodos actuales.
- Profesional encargado de diseño, desarrollo e implementación de las entrevistas a personas que cumplan con los requisitos.
- Profesional encargado de análisis, diseño, desarrollo e integración del sistema de monitoreo.
- Investigación del funcionamiento y aplicación de los algoritmos de redes neuronales para el reconocimiento de imágenes.
- Acceso a información documental de los algoritmos de inteligencia artificial para su implementación en el prototipo.

- Acceso a documentación técnica de los *frameworks* y lenguajes de programación utilizados para la implementación de redes neuronales.
- Acceso a librerías para alimentación de los algoritmos de inteligencia artificial.
- Programación e implementación de los algoritmos de inteligencia artificial a utilizar.
- Creación de aplicación móvil para la gestión de alertas al usuario final.
- Construcción de un sistema de videovigilancia en un entorno cerrado mediante el uso de cámaras IP y red local para el monitoreo.

El objetivo del proyecto es automatizar el monitoreo para demostrar que los métodos de inteligencia artificial optimizan el proceso y lo realizan de manera eficiente. Para ello se hace el estudio del método convencional eligiendo a una población que tenga los requisitos mínimos para poder ejecutar el estudio. Para el desarrollo del prototipo se deberá realizar una investigación, diseño y desarrollo del sistema donde estén involucrados las redes neuronales y el sistema de monitoreo.

Se puede concluir que las personas que participen en el estudio pueden ser de diferentes ámbitos, debido a que lo importante es la observación en el sistema, por lo que es factible conseguir personal para realizarlo y lograr la investigación de los métodos actuales y elegirlos. Existen varios artículos científicos, tesis, cursos en los que se puede obtener información acerca de los métodos por lo que es posible acceder a esta información, conocer su implementación para desarrollar el prototipo y para los componentes electrónicos

necesarios se posee la información técnica necesaria en los manuales de los componentes. Para el desarrollo del prototipo; el encargado de supervisar estos eventos será el investigador debido a que cuenta con lo requerido para implementarlo, por lo que es factible el desarrollo del prototipo desde el punto de vista operativo.

## **12.2. Factibilidad técnica**

Para el desarrollo y ejecución del prototipo se necesitará de los siguientes recursos técnicos.

- Profesional con experiencia en desarrollo de aplicaciones, base de datos, servicios en la nube, redes y conocimientos de inteligencia artificial.
- El tiempo de desarrollo se estable como 8 meses para lograr el desarrollo y análisis de los resultados por ello se distribuirá el tiempo, trabajando 3 horas diarias de lunes a viernes.
- Equipo de computación que cumpla con los requisitos mínimos:
  - Procesador i7 9na generación
  - 16GB de RAM
  - Disco duro 500GB
  - Tarjeta gráfica: NVIDIA con 6GB de RAM
  - Equipo portátil
- Herramientas necesarias para el desarrollo del prototipo
  - *Framework*

- Base de datos
- IDE
- Servicio en la nube para alojar el servidor que permita el reconocimiento de imágenes.
- Redes inalámbricas/alámbricas para la conexión de la red con las cámaras IP.
- Cámaras IP que cumplan con los requerimientos mínimos:
  - 2MP
  - Resolución HD1080P
  - Conexión WIFI/Ethernet

En la información de los requerimientos técnicos para el desarrollo del prototipo se necesita de un profesional que cumpla con el perfil de un ingeniero en ciencias y sistemas, por lo que el investigador tomará este rol para el desarrollo del sistema, el cual cuenta con el conocimiento de herramientas que pueden ser utilizadas para el desarrollo y conocimiento de servicios en la nube y redes. Se necesitarán recursos como las cámaras y el equipo; el investigador los proporcionará.

### **12.3. Factibilidad económica**

Para la ejecución del prototipo se necesitarán los recursos que se presentan en la tabla detallando su uso y el costo monetario.

Tabla II. Presupuesto

Recurso	Descripción	Costo
<b>Profesional</b>	Diseño, desarrollo e implementación de módulo de reconocimiento, aplicación móvil y sistema externo de videovigilancia. Costo por hora de Q.130. Tiempo de trabajo: 15 horas de trabajo a la semana durante 8 meses.	Q. 62,400.00
<b>Asesoría profesional</b>	Asesoría por trabajo de graduación.	Q. 2,500.00
<b>Equipo de computo</b>	Computadora para desarrollo del sistema y pruebas de ejecución <ul style="list-style-type: none"> <li>• Procesador i7 9na generación</li> <li>• 16GB de RAM</li> <li>• Disco Duro 500GB</li> <li>• Tarjeta Gráfica: NVIDIA con 6GB de RAM</li> <li>• Equipo portátil</li> </ul>	Q.15,000.00
<b>Sistema de videovigilancia</b>	Cámaras IP, conexiones de red y otros dispositivos. Requerimiento de las cámaras <ul style="list-style-type: none"> <li>• 2MP</li> <li>• Resolución HD1080P</li> <li>• Conexión WIFI/Ethernet</li> </ul>	Q.3,500.00
<b>Infraestructura tecnológica en la nube</b>	Servicios de AWS para el servidor del módulo de reconocimiento y uso de base de datos que permite el adecuado manejo de la información en el sistema <ul style="list-style-type: none"> <li>• EC2 a costo de Q.350 por mes</li> <li>• DynamoDB (se parte de la capa gratuita)</li> </ul>	Q.3,500.00
<b>Gastos fijos</b>	Costo de servicios generales como agua, energía eléctrica, internet, entre otros. Costo mensual de Q.650 por mes.	Q.5,200.00
<b>Total</b>		Q.92,600.00

Fuente: elaboración propia, realizado con Word

El costo del prototipo es de Q 92,600.00, en este caso el mayor monto está concentrado en lo que son los honorarios del profesional, entonces debido a que este papel será ejecutado por el investigador y también será el encargado de proporcionar el equipo de cómputo, al restar los montos respectivos queda un resto de Q 14,000.00 lo cual será cubierto por el investigador, por lo que se concluye que el estudio es factible económicamente debido a que se posee la capacidad de cubrir el presupuesto planteado.



### 13. REFERENCIAS

1. Fernández, J. R. (2013). *Circuito cerrado de televisión y seguridad electrónica*. España: Ediciones Paraninfo, SA.
2. Flórez López, R., & Fernández Fernández, J. M. (2008). *Las Redes Neuronales Artificiales, Fundamentos teoricos y aplicaciones practicas*. España: Netbiblo, S.L.
3. Jimenez, R. A. (2010). *EVENT DETECTION IN SURVEILLANCE VIDEO*. Florida.
4. LI, L., & Haibin, X. (2019). A New Intelligent Video Surveillance Architecture. *Journal of Physics: Conf. Serie 1168*.
5. Mata, F. J. (2010). *Videovigilancia: CCTV usando video IP*. España: Publicaciones Vértice S. L.
6. Osorio, R., Olivera, W., Peña, M., Lopez Juárez, I., & Lefranc, G. (2016). *Video Vigilancia Utilizando Algoritmo de Segmentación*. Departamento de Ingeniería de Sistemas Computacionales y Automatización. Instituto de Investigaciones en Matemáticas.
7. Pérez Borrero, I., & Gegúndez Arias, M. E. (2021). *Deep Learning Fundamentos, teoria y aplicacion*. España: Publicaciones de la Universidad de Huelva.

8. Pérez, E. H. (2010). *TECNOLOGÍAS Y REDES DE TRANSMISIÓN DE DATOS*. México: LIMUSA S.A.
9. Rouhiainen, L. (2018). *Inteligencia Artificial*. España: Alienta.
10. Salgado, Y. A. (2016). *Detección de cuchillos con cámaras de video vigilancia en interiores*. México: Tecnológico Nacional de México, Instituto Tecnológico de Celaya.
11. Secretaría de estado de digitalización e inteligencia artificial. (28 de 01 de 2020). *Tecnologías Emergentes y Datos Abiertos* Obtenido de [https://datos.gob.es/sites/default/files/doc/file/tecnologias\\_emergentes\\_y\\_opendata\\_inteligencia\\_artificial-\\_final\\_0.pdf](https://datos.gob.es/sites/default/files/doc/file/tecnologias_emergentes_y_opendata_inteligencia_artificial-_final_0.pdf).
12. Torres, J. (2018). *Deep Learning Introducción práctica con Keras*. Barcelona: WHAT THIS SPACE.
13. Verma, G., & Tiwari, R. (2015). A Computer Vision based Framework for Visual Gun Detection Using Harris Interest Point Detector. *Conference: Eleventh International Conference on Communication Networks* (págs. 21-23). Bangalore, India: Bangalore.
14. Vilet, J. R. (2005). *Apuntes de Procesamiento Digital de Imágenes*. Obtenido de <http://laurence.com.ar/artes/comun/Apuntes%20procesamiento%20digital%20de%20imagenes.pdf>.

15. Zúñiga Ramos, D. A., Dongo Esquivel, D. Y., Gonzáles Manrique, L., Naola Pereyra, Y. A., & Ttito, E. L. (2019). *Reconocimiento de objetos mediante Visión Artificial*. Escuela Profesional de Ingeniería Informática.