**UNIVERSIDAD MAYOR REAL Y PONTIFICIA DE SAN FRANCISCO XAVIER DE CHUQUISACA**

**FACULTAD DE CIENCIA Y TECNOLOGÍA**



**Detección y Seguimiento Del Personal**

**Materia**: Desarrollo de aplicaciones inteligentes (SIS-330)

**Docente**: Ing. Carlos Walter Pacheco Lora

**Universitario**: Choquerive Quispe Erwing

**Carrera**: Ing. Ciencias de la computación

**Fecha**: 25-09-2024

Contenido

[Antecedentes 3](#_Toc178157676)

[Problemática 3](#_Toc178157677)

[Abordaje o Propuesta De Solución 3](#_Toc178157678)

[Objetivo General 3](#_Toc178157679)

[Objetivos Específicos 3](#_Toc178157680)

[Técnicas De Recolección De Datos 4](#_Toc178157682)

[Materiales y Herramientas 4](#_Toc178157683)

[Etapas De Trabajo 5](#_Toc178157684)

[Cronograma De Proyecto 5](#_Toc178157685)

[Resultados y Conclusiones 6](#_Toc178157686)

# Antecedentes

El registro de ingreso, la búsqueda y el seguimiento del personal en universidades, edificios y lugares de trabajo son esenciales para garantizar la asistencia y el cumplimiento de las labores. Sin embargo, esto no asegura que el personal esté realizando las actividades asignadas durante su horario de trabajo, lo que requiere la intervención de supervisores para revisar dichas actividades. Este control es fundamental para verificar la asistencia, asegurar el cumplimiento de las tareas asignadas y supervisar las diversas actividades que se llevan a cabo en estos entornos.

# Problemática

La necesidad de identificar al personal y determinar su ubicación en tiempo real mediante cámaras de seguridad presenta varias limitaciones. Los sistemas tradicionales dependen en gran medida de la intervención humana, lo que puede generar errores, retrasos en la identificación y una falta de precisión en el monitoreo continuo.

# Abordaje o Propuesta De Solución

Para resolver este problema, se propone desarrollar un software web con inteligencia artificial basado en visión por computadora, que permita la detección y seguimiento del personal en tiempo real a través de las cámaras de seguridad. Este software integrará algoritmos de reconocimiento facial y de identificación de características como la vestimenta del personal. Para su implementación, se utilizarán modelos que detecten el rostro, la persona y la ropa, y, mediante redes siamesas, se logrará la identificación precisa de los individuos a través de las videocámaras.

# Objetivo General

Desarrollar un software web con inteligencia artificial capaz de identificar y seguir automáticamente a personas específicas a través de múltiples cámaras de seguridad, mejorando la eficacia y precisión del monitoreo en tiempo real, con un objetivo de precisión superior al 75%.

# Objetivos Específicos

* Reducir la dependencia de la intervención humana en la vigilancia, mejorando la eficiencia y precisión del monitoreo.
* Implementar técnicas avanzadas de detección y seguimiento de personas mediante inteligencia artificial.
* Desarrollar un software web que integre múltiples cámaras y permita un seguimiento continuo de individuos específicos.
* Realizar pruebas reales para verificar la precisión y efectividad de los resultados

# Técnicas De Recolección De Datos

Se realizó la búsqueda de datasets ya preparados o listos para entrenar:

Dataset para reconocimiento de personas: <https://github.com/SMSajadi99/Custom-Data-YOLOv8-Person-Detection/tree/main>

Dataset usado para el entrenamiento de reconocimiento de ropa: <https://universe.roboflow.com/dataset-yn4f8/main-fashion/dataset/6>

Detección de rostro: Se realizan capturas fotográficas que contienen la cara, y el mismo proyecto de reconocimiento facial se encarga de obtener las caras detectadas. <https://github.com/vectornguyen76/face-recognition/tree/master>

# Materiales y Herramientas

* + **Redes siamesas**: Las redes siamesas son un tipo de arquitectura de red neuronal que utiliza dos o más subredes idénticas que comparten los mismos pesos y parámetros. Se utilizan comúnmente para comparar similitudes entre dos entradas, como en la verificación de identidad facial, donde se determina si dos imágenes son de la misma persona o no.
  + **Yolo v5 face**: Es una versión especializada del modelo YOLO v5 (You Only Look Once) adaptada para la detección de rostros. YOLOv5 es un modelo de detección de objetos en tiempo real que ofrece alta precisión y velocidad, y la variante de detección de caras se entrena específicamente para localizar y clasificar rostros en imágenes o videos.
  + **Yolo v8**: La versión 2023 de la familia YOLO (You Only Look Once), diseñada para la detección de objetos en tiempo real. YOLOv8 incluye mejoras en velocidad, precisión y facilidad de uso en comparación con sus predecesores, utilizando arquitecturas más eficientes y técnicas avanzadas de entrenamiento.
  + **Flask**: Es un microframework web escrito en Python. Flask es ligero y flexible, diseñado para facilitar el desarrollo rápido de aplicaciones web, permitiendo una mayor personalización y control sobre componentes como rutas y bases de datos, sin imponer un gran número de restricciones.
  + **Pytorch**: Es una biblioteca de código abierto para el aprendizaje profundo y la computación numérica. PyTorch es ampliamente utilizado para desarrollar y entrenar redes neuronales gracias a su facilidad de uso y su capacidad para realizar cálculos automáticos en tensores, ofreciendo un enfoque dinámico para la creación de modelos de IA.
  + **Python 3.9**: Es el lenguaje de programación utilizado como base para el funcionamiento y proceso del proyecto, se hace uso de la versión 3.9 de Python
  + **ResNet18**: Es una arquitectura de red neuronal profunda de la familia ResNet (Redes Residuales), diseñada para la clasificación de imágenes. ResNet18 tiene 18 capas y se distingue por utilizar conexiones residuales ("skip connections") que permiten entrenar redes más profundas sin el problema de la degradación, mejorando la precisión en tareas de visión por computadora.

# Etapas De Trabajo

* + Realizar sobre redes siamesas y búsqueda inversa de imágenes.
  + Buscar un modelo básico en formato .pt que permita la detección de personas.
  + Realizar pruebas de búsqueda inversa utilizando un video y una imagen de referencia.
  + Realizar pruebas con redes siamesas utilizando un video y una imagen de referencia.
  + Integrar el modelo preentrenado ResNet18 en las redes siamesas.
  + Mejorar e integrar el modelo de detección de personas.
  + Integrar un modelo de detección de ropa en las redes siamesas.
  + Implementar mejoras para aumentar la precisión en la detección de personas: identificación del color de la ropa, validaciones y recorte de la imagen de entrada.
  + Preparar el entorno del proyecto con detección de rostros.
  + Desarrollar la parte de diseño frontal para la interacción con el usuario.
  + Identificar e implementar mejoras necesarias.

# Cronograma De Proyecto

|  |  |
| --- | --- |
| Etapa | Duración |
| Pruebas básicas con búsqueda inversa y redes siamesas a partir de un video y una imagen de referencia, con el objetivo de analizar el funcionamiento de ambas redes y determinar cuál ofrece mejores resultados. | Hasta el 28 de agosto |

|  |  |
| --- | --- |
| Búsqueda de un modelo para detectar personas y su integración con redes siamesas para realizar pruebas utilizando una foto a través de la videocámara, además de la integración de ResNet18. | Del 28 al 04 de septiembre |
| Búsqueda de un modelo para la detección de ropa y su integración en redes siamesas para mejorar la precisión en la identificación del individuo buscado. | Del 04 al 11 de septiembre |
| Exploración de mejoras, como la obtención del color de la ropa y la implementación de validaciones adicionales, con el fin de aumentar la precisión de la búsqueda | Del 11 al 18 de septiembre |
| Desarrollo de la interfaz de usuario basada en un proyecto de reconocimiento facial, incluyendo la integración visual con Ninja y la comunicación con el servidor mediante Flask y Sockets para el envío de fotogramas. | Del 18 al 25 de septiembre |

# Resultados y Conclusiones

Las pruebas realizadas muestran que, con las validaciones actuales, el sistema solo es capaz de detectar al individuo cuando este mantiene una pose similar a la de la imagen de referencia utilizada para la detección. Además, se observó que si el individuo se encuentra muy alejado de la cámara, el modelo no logra identificarlo de manera precisa. Asimismo, presenta dificultades para detectar al individuo cuando está de espaldas o de frente, o cuando comparte características similares con otras personas cercanas.

En conclusión, el modelo presenta fallas importantes en la detección, especialmente en situaciones de variabilidad de poses, distancias y características compartidas entre individuos. Las validaciones actuales son insuficientes y requieren ser mejoradas para incrementar la precisión del sistema. Para mejorar el rendimiento, es necesario integrar nuevas técnicas de validación y reconocimiento que sean más robustas frente a variaciones en poses, ángulos y características físicas, así como optimizar la detección a diferentes distancias.