



Sistema de Re-Identificación de personas a través de las características biométricas: facial y la forma de caminar y softbiométricas: silueta corporal, textura y color de la vestimenta, con el uso de modelos no holísticos, aplicando técnicas computacionales avanzadas.

Chicaiza Chicaiza, Bryan Stiven y Arteaga Chapi, Lenin Sebastián

Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

Trabajo de integración curricular, previo a la obtención del título de Ingeniero en
Software

Dr. Carillo Medina, José Luis

18 de agosto del 2023

Reporte de verificación de contenido



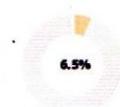
Plagiarism and AI Content Detection Report

Tesina_capitulos_Chicaiza-Arteaga_20...

Scan details

Scan time: August 22th, 2023 at 23:42 UTC
 Total Pages: 114
 Total Words: 28276

Plagiarism Detection



Types of plagiarism	Words
Identical	0.4% 111
Minor Changes	0.1% 34
Paraphrased	4.8% 1360
Omitted Words	17.7% 5014

AI Content Detection



Text coverage	Words
AI text	5.9% 594
Human text	94.1% 25679

[Learn more](#)

Alerts: (1)

Cross Language: Same Document Language

Submitted language and cross-language text are the same language. No credits were used.

2/5 Severity



Plagiarism Results: (34)

CARACTERIZACIÓN MACROPROCESO GESTIÓN FINA... 0.9%

<https://repositorio.espe.edu.ec/bitstream/21000/35752/2/es...>

ESPE

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN CARRERA DE INGENIERÍA DE SOFTWARE TRABAJO DE UNIDAD DE INTEGRACIÓN CURRICULAR PREVIO ...

Redalyc.Re-identificación de personas a través de s... 0.8%

<https://www.redalyc.org/pdf/404/40445803010.pdf>

Moctezuma-Ochoa, Daniela Alejandra

Ingeniería. Investigación y Tecnología ISSN: 1405-7743 iit.revista@gmail.com Universidad Nacional Autónoma de México México Moctezuma-Oc...

Certified by

About this report
help.copyleaks.com

copyleaks.com

Firma:

Ing. Carrillo Medina, José Luis, Ph.D

C. C: 0501553788



Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

Certificación

Certifico que el trabajo de integración curricular: **“Sistema de Re-Identificación de personas a través de las características biométricas: facial y la forma de caminar y softbiométricas: silueta corporal, textura y color de la vestimenta, con el uso de modelos no holísticos, aplicando técnicas computacionales avanzadas”** fue realizado por los señores **Chicaiza Chicaiza, Bryan Stiven y Arteaga Chapi, Lenin Sebastián**; el mismo que cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, además fue revisado y analizado en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

Latacunga, 23 de agosto del 2023

Firma:

Ing. Carrillo Medina, José Luis, Ph.D

C. C: 0501553788



Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

Responsabilidad de Autoría

Nosotros, **Chicaiza Chicaiza, Bryan Stiven y Arteaga Chapi, Lenin Sebastián**, con cédulas de ciudadanía n° 0503680266 y n° 1720409265, declaramos que el contenido, ideas y criterios del trabajo de integración curricular: **“Sistema de Re-Identificación de personas a través de las características biométricas: facial y la forma de caminar y softbiométricas: silueta corporal, textura y color de la vestimenta, con el uso de modelos no holísticos, aplicando técnicas computacionales avanzadas.”** es de nuestra autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Latacunga, 23 de agosto del 2023

Firma

Chicaiza Chicaiza, Bryan Stiven

C.C.: 0503680266

Firma

Arteaga Chapi, Lenin Sebastián

C.C.: 1720409265



Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

Autorización de Publicación

Nosotros **Chicaiza Chicaiza, Bryan Stiven y Arteaga Chapi, Lenin Sebastián**, con cédulas de ciudadanía n° 0503680266 y n° 1720409265, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de integración curricular: **“Sistema de Re-Identificación de personas a través de las características biométricas: facial y la forma de caminar y softbiométricas: silueta corporal, textura y color de la vestimenta, con el uso de modelos no holísticos, aplicando técnicas computacionales avanzadas.”** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi/nuestra responsabilidad.

Latacunga, 23 de agosto del 2023

Firma


.....
Chicaiza Chicaiza, Bryan Stiven

C.C.: 0503680266

Firma


.....
Arteaga Chapi, Lenin Sebastián

C.C.: 1720409265

Dedicatoria

A mis padres y hermanos, quienes han sido mi apoyo y motivación a lo largo de proceso de integración curricular. Dedicado a todos los que fueron de apoyo en a lo largo de este trabajo.

A mi tutor de proyecto, Ing. José Carrillo, quien con su paciencia, apoyo, sugerencias e ideas me ayudaron a mejorar mi trabajo y mantenerme motivado en esta área de investigación.

Bryan Stiven Chicaiza Chicaiza

Ecuador, agosto 2023

Dedicatoria

A mi padre Cesar Arteaga, el cual siempre estuvo a mi lado a lo largo de toda mi vida. Gracias por apoyarme incluso cuando ni siquiera yo mismo creía en mí.

Lenin Sebastián Arteaga Chapi

Ecuador, agosto 2023

Agradecimiento

Agradezco a mi madre Marina Chicaiza Vázquez quien me brindo su apoyo a lo largo de toda la carrera, manteniéndome motivado y mantener la perseverancia. Sus consejos y palabras de ánimo me ayudaron a no rendirme durante los tiempos difíciles.

Agradezco a todos quienes me guiaron y colaboraron en este trabajo, sin su tiempo y esfuerzo este trabajo no habría sido posible.

Bryan Stiven Chicaiza Chicaiza

Ecuador, agosto 2023

Agradecimiento

A todas las personas con las cuales en algún momento cruce mi camino y me supieron extender una mano, incluso aquellos de los cuales no esperaba recibir ayuda.

A todos ellos gracias por enseñarme que no todo es blanco y negro y sobre todo porque con su ayuda desinteresada aprendí a ser un poco más humilde.

Lenin Sebastián Arteaga Chapi

Ecuador, agosto 2023

ÍNDICE DE CONTENIDO

Carátula	1
Reporte de verificación de contenido.....	2
Certificación	3
Responsabilidad de autoría	4
Autorización de Publicación	5
Dedicatoria	6
Dedicatoria	7
Agradecimiento.....	8
Agradecimiento.....	9
Índice de contenido.....	10
Índice de figuras.....	14
Índice de tablas	17
Resumen	19
Abstract	20
Capítulo I: Introducción.....	21
Propósito y contextualización del tema	21
Justificación del interés de la investigación	21
Objetivos Generales y Específicos	22
<i>Objetivo General</i>	23
<i>Objetivos Específicos</i>	23
Metodología	23
Capítulo II: Marco Teórico	28
Introducción al Marco Teórico	28

Características Biométricas y Soft-biométricas	28
<i>Características Biométricas</i>	29
<i>Características Soft-biométricas</i>	30
Importancia en la Re-Identificación	32
Técnicas Holísticas y No Holísticas	32
<i>Modelos Holísticos</i>	33
<i>Modelos No Holísticos</i>	34
Técnicas de Visión por Computadora	35
Local Binary Pattern (LBP).....	37
Algoritmo de Viola-Jones	39
Support Vector Machine (SVM)	42
Redes Neuronales Convolucionales (CNN)	45
Gait Energy Image (GEI)	47
Histogram of Oriented Gradients (HoG).....	48
You Only Look Once (YOLO).....	50
Proceso de Re-Identificación y Fusión de Modelos	51
Re-Identificación facial y forma de caminar	51
Re-Identificación basada en color, textura y silueta corporal.....	53
Proceso de fusión de características y modelos	54
Tecnologías de implementación	55
Métricas de evaluación	56
Marco de trabajo	59
Capítulo III: Proceso de implementación del sistema de Re-Identificación de personas	61
Descripción del equipo Scrum.....	61

Historias de Usuario.....	62
Product Backlog del proyecto	67
Ambiente para la implementación del sistema de Re-Identificación de personas.....	69
Sprint 1: Inicio y Preparación	71
<i>Historia de Usuario 01: Creación de Dataset que contenga características biométricas y softbiométricas.</i>	72
<i>Sprint Backlog 01</i>	73
<i>Burndown chart sprint 01</i>	74
<i>Resultados Sprint 01</i>	76
Sprint 2: Desarrollo de modelo facial.....	76
<i>Historia de usuario 02: Desarrollo de modelo facial.</i>	78
<i>Sprint Backlog 02</i>	79
<i>Burndown chart sprint 02</i>	81
<i>Resultados Sprint 02</i>	82
Sprint 3: Desarrollo de modelo para la forma de caminar.....	87
<i>Historia de usuario 03: Desarrollo de modelo para la forma de caminar</i>	88
<i>Sprint Backlog 03</i>	90
<i>Burndown chart sprint 03</i>	93
<i>Resultados Sprint 03</i>	94
Sprint 4: Desarrollo de Modelos Soft-biométricos	98
<i>Historia de Usuario 04: Desarrollo de Modelos Soft-biométricos</i>	101
<i>Sprint Backlog 04</i>	102
<i>Burndown chart sprint 04</i>	104
<i>Resultados Sprint 04</i>	105

Sprint 5: Fusión de Modelos	110
<i>Historia de Usuario 05: Fusión de Modelos.....</i>	<i>111</i>
<i>Sprint Backlog 05.....</i>	<i>112</i>
<i>Burndown chart sprint 05.....</i>	<i>114</i>
<i>Resultados Sprint 05.....</i>	<i>114</i>
Sprint 6: Validación, Ajustes e Implantación	117
<i>Historia de Usuario 06: Validación, Ajustes e Implantación.....</i>	<i>118</i>
<i>Sprint Backlog 06.....</i>	<i>119</i>
<i>Burndown chart sprint 06.....</i>	<i>120</i>
<i>Resultados Sprint 06.....</i>	<i>121</i>
Capítulo IV: Validación del sistema	124
Marco de Evaluación.....	124
Aplicación de métricas de evaluación	125
Pruebas del algoritmo de identificación facial	125
Pruebas del algoritmo de Re-Identificación de color, textura y silueta	128
Pruebas del algoritmo de Re-Identificación forma de andar	131
Capítulo V: Conclusiones	140
Bibliografía.....	143
Anexos	151

ÍNDICE DE FIGURAS

Figura 1 <i>Ciclo de la caminata</i>	29
Figura 2 <i>Características soft-biométricas</i>	31
Figura 3 <i>Representación de modelo holístico</i>	33
Figura 4 <i>Representación de modelo no holístico</i>	34
Figura 5 <i>Asignación de valores a píxeles mediante LBP</i>	37
Figura 6 <i>Selección de regiones de una imagen facial</i>	38
Figura 7 <i>Características de Haar</i>	39
Figura 8 <i>Regiones identificadas por el clasificador Adaboost</i>	40
Figura 9 <i>Representación de SVM</i>	42
Figura 10 <i>Clasificación en una red de vectores de soporte (SVM)</i>	43
Figura 11 <i>Proceso de una red neuronal convolucional (CNN)</i>	46
Figura 12 <i>Representación de siluetas en una secuencia de caminata</i>	47
Figura 13 <i>Descriptor de características HOG</i>	49
Figura 14 <i>Representación de YOLO</i>	50
Figura 15 <i>Ciclo de vida de Scrum</i>	60
Figura 16 <i>Entorno controlado para pruebas del sistema de Re-Identificación de personas</i>	70
Figura 17 <i>Ubicación de las cámaras utilizadas para capturar los videos de prueba</i>	70
Figura 18 <i>Burndown chart Sprint 01</i>	75
Figura 19 <i>Videos capturados para la generación automática de datasets</i>	76
Figura 20 <i>Descripción general del sistema de reconocimiento facial</i>	77
Figura 21 <i>Burndown chart 02</i>	81
Figura 22 <i>Pseudocódigo del algoritmo utilizado para la identificación del rostro</i>	82
Figura 23 <i>Segmento de código utilizado para detectar el rostro y generar el dataset</i>	83

Figura 24 <i>Imágenes generadas para el entrenamiento del algoritmo del rostro</i>	83
Figura 25 <i>Segmento de código utilizado para el entrenar el algoritmo del rostro</i>	84
Figura 26 <i>Modelo resultante del entrenamiento del algoritmo de identificación del rostro</i>	84
Figura 27 <i>Código de identificación de un individuo en los videos que simulan las cámaras 1 y 2</i>	85
Figura 28 <i>Identificación facial de una persona</i>	85
Figura 29 <i>Matriz de confusión del resultado de identificación facial de la clase Danilo</i>	86
Figura 30 <i>Descripción general del sistema de Re-Identificación por la forma de andar (cámara 1)</i>	88
Figura 31 <i>Descripción general del sistema de Re-Identificación por la forma de andar (cámara 2)</i>	89
Figura 32 <i>Burndown chart sprint 03</i>	93
Figura 33 <i>Pseudocódigo del algoritmo utilizado para la identificación de la forma de andar</i>	94
Figura 34 <i>Código utilizado para detectar una persona, obtener su silueta y generar el dataset</i>	94
Figura 35 <i>Imágenes generadas para el entrenamiento del modelo de la forma de andar</i>	95
Figura 36 <i>Segmento de código utilizado para entrenar el modelo de la forma de andar</i>	95
Figura 37 <i>Modelo resultante de Re-Identificación por la forma de andar</i>	96
Figura 38 <i>Matriz de confusión de Re-Identificación por la forma de andar de la clase Danilo</i>	96
Figura 39 <i>Identificación facial de una persona por la forma de andar, cámara 1</i>	98
Figura 40 <i>Identificación facial de una persona por la forma de andar, cámara 2</i>	98
Figura 41 <i>Esquema de Re-Identificación por características Soft-biométricas (cámara 1)</i>	99
Figura 42 <i>Esquema de Re-Identificación por características Soft-biométricas (cámara 2)</i>	100
Figura 43 <i>Burndown chart sprint 04</i>	104
Figura 44 <i>Pseudocódigo del algoritmo para la identificación softbiométrica</i>	105
Figura 45 <i>Código utilizado para detectar silueta corporal, color y textura de la vestimenta</i>	106
Figura 46 <i>Imágenes generadas para el entrenamiento del modelo Soft-biométrico</i>	107
Figura 47 <i>Segmento de código utilizado para entrenar el modelo Soft-biométrico</i>	107

Figura 48 <i>Modelo de Re-Identificación por características soft-biométricas</i>	107
Figura 49 <i>Código utilizado para la Re-Identificación por características Soft-biométricas</i>	108
Figura 50 <i>Matriz de confusión de Re-Identificación por características Soft-biométricas</i>	108
Figura 51 <i>Identificación de una persona por características soft-biométricas (cámara 1)</i>	110
Figura 52 <i>Identificación de una persona por características soft-biométricas (cámara 2)</i>	110
Figura 53 <i>Burndown chart sprint 05</i>	114
Figura 54 <i>Pseudocódigo del algoritmo de fusión mediante promedio ponderados</i>	115
Figura 55 <i>Código para realizar la fusión de características a través del promedio ponderado</i>	115
Figura 56 <i>Identificación de una persona con la fusión de los tres modelos generados, cámara 1</i>	116
Figura 57 <i>Identificación de una persona con la fusión de los tres modelos generados, cámara 2</i>	116
Figura 58 <i>Burndown chart sprint 06</i>	121
Figura 59 <i>Arquitectura del modelo CNN para la Re-Identificación de personas</i>	121
Figura 60 <i>Aumento de datos para mejorar la precisión del modelo por la forma de andar</i>	122
Figura 61 <i>Primera arquitectura CNN propuesta para la Re-Identificación por la forma de andar</i>	133
Figura 62 <i>Primer modelo utilizado para la Re-Identificación de personas por la forma de andar</i>	133
Figura 63 <i>Arquitectura CNN modificada, para la Re-Identificación por la forma de andar</i>	137
Figura 64 <i>Resultados del nuevo modelo utilizado para la Re-Identificación por la forma de andar</i>	138

ÍNDICE DE TABLAS

Tabla 1 <i>Métricas de evaluación utilizadas</i>	57
Tabla 2 <i>Valores de la matriz de confusión</i>	58
Tabla 3 <i>Equipo Scrum</i>	61
Tabla 4 <i>Historias de Usuario</i>	62
Tabla 5 <i>Product Backlog del Proyecto</i>	68
Tabla 6 <i>Distribución de los Sprint del proyecto</i>	68
Tabla 7 <i>Historia de usuario para la obtención del dataset</i>	72
Tabla 8 <i>Sprint Backlog 01</i>	73
Tabla 9 <i>Historia de usuario 02: Desarrollo de modelo facial</i>	78
Tabla 10 <i>Sprint Backlog 02</i>	79
Tabla 11 <i>Resultados del algoritmo de identificación de personas por el rostro</i>	87
Tabla 12 <i>Historia de usuario para el modelo de la forma de caminar.</i>	90
Tabla 13 <i>Sprint Backlog 03</i>	91
Tabla 14 <i>Código utilizado para la Re-Identificación por la forma de andar</i>	96
Tabla 15 <i>Resultados del algoritmo de Re-Identificación por la forma de andar de la clase Danilo</i>	97
Tabla 16 <i>Historia de usuario para el modelo Soft-biométrico</i>	101
Tabla 17 <i>Sprint Backlog 04</i>	102
Tabla 18 <i>Resultados del algoritmo de Re-Identificación por características Soft-biométricas</i>	109
Tabla 19 <i>Historia de usuario para fusión de modelos</i>	111
Tabla 20 <i>Sprint Backlog 05</i>	112
Tabla 21 <i>Historia de usuario para validación, ajustes e implantación</i>	118
Tabla 22 <i>Sprint Backlog 06</i>	119
Tabla 23 <i>Resultado de acuraccy del primer y segundo modelo obtenidos al utilizar la cámara 2</i>	122

Tabla 24 <i>Matriz de confusión del modelo de identificación facial de la clase Danilo</i>	125
Tabla 25 <i>Matriz de confusión del modelo de identificación facial de la clase Luis</i>	126
Tabla 26 <i>Matriz de confusión del modelo de identificación facial de la clase Pedro</i>	126
Tabla 27 <i>Resultados de la exactitud del modelo de las 3 clases mediante la identificación facial</i>	127
Tabla 28 <i>Porcentaje de predicción del modelo de identificación facial</i>	128
Tabla 29 <i>Matriz de confusión del modelo combinado de la clase Danilo</i>	129
Tabla 30 <i>Matriz de confusión del modelo combinado de la clase Luis</i>	129
Tabla 31 <i>Matriz de confusión del modelo combinado de la clase Pedro</i>	129
Tabla 32 <i>Resultados de exactitud del modelo Soft-biométrico combinado</i>	130
Tabla 33 <i>Porcentaje de predicción del modelo soft-biométrico combinado</i>	130
Tabla 34 <i>Matriz de confusión del modelo de la forma de andar de la clase Danilo</i>	131
Tabla 35 <i>Matriz de confusión del modelo de la forma de andar de la clase Luis</i>	132
Tabla 36 <i>Matriz de confusión del modelo de la forma de andar de la clase Pedro</i>	132
Tabla 37 <i>Resultados de la exactitud del modelo combinado</i>	134
Tabla 38 <i>Porcentaje de predicción del modelo de la forma de andar</i>	135
Tabla 39 <i>Matriz de confusión del nuevo modelo de la forma de caminar de la clase Danilo</i>	136
Tabla 40 <i>Matriz de confusión del nuevo modelo de la forma de caminar combinado de la clase Luis</i>	136
Tabla 41 <i>Matriz de confusión del nuevo modelo de la forma de caminar de la clase Pedro</i>	136
Tabla 42 <i>Resultados de la exactitud del modelo combinado</i>	138
Tabla 43 <i>Porcentaje de predicción del modelo de la forma de caminar</i>	139

Resumen

El presente trabajo tiene como objetivo desarrollar un sistema de Re-Identificación de personas mediante el uso de rasgos soft-biométricos (silueta corporal, color y textura de la ropa) y rasgos biométricos (rostro y marcha), utilizando modelos no holísticos, técnicas de Visión por Computadora y modelos y/o algoritmos de Machine Learning. Para ello se planteó un escenario controlado el cual se encuentra ubicado en los pasillos de la Universidad de las Fuerzas Armadas ESPE Sede Latacunga. Este escenario consta de dos cámaras de videovigilancia, donde la primera cámara extrae las características biométricas y soft-biométricas de una persona mientras esta se encuentra caminando frente a la cámara, estos datos permiten a los modelos desarrollados obtener patrones únicos que servirán para un posterior reconocimiento. A su vez, la segunda cámara de seguridad re-identifica a la persona mientras esta se encuentra de espaldas. Este sistema se desarrolló a través de un enfoque metodológico ágil por medio de la metodología Scrum, donde se llevaron a cabo 6 Sprints, cada uno con objetivos y entregables específicos. El primer sprint se centró en la obtención y preparación de datasets adecuados. Los sprints subsiguientes se dedicaron al desarrollo de modelos individuales: reconocimiento facial, identificación por la forma de caminar y características soft-biométricas. Posteriormente, se trabajó en la fusión de estos modelos para crear un sistema robusto y preciso. Finalmente, se realizó una fase de validación, ajustes y documentación. El sistema propuesto demostró ser eficiente en pruebas preliminares, presentado resultados acordes a la revisión de la literatura del proyecto.

Palabras clave: Re-Identificación de personas, características biométricas, características soft-biométricas, visión por computadora, machine learning, scrum.

Abstract

This research work aims to develop and implement a system of Re-Identification of people through the use of biometric characteristics (face and gait) and soft-biometric characteristics (body silhouette, color and texture of clothing) using non-holistic models, techniques, models and / or computer vision algorithms and machine learning. This scenario consists of two video surveillance cameras, where the first camera extracts the biometric and soft-biometric characteristics of a person while he/she is walking in front of the camera, this data allows the developed models to obtain unique patterns that will be used for later recognition. At the same time, the second security camera re-identifies the person while he/she is on his/her back. This system was developed through an agile methodological approach using the Scrum methodology, where six sprints were carried out, each with specific objectives and deliverables. The first sprint focused on obtaining and preparing suitable datasets. Subsequent sprints were devoted to the development of individual models: face recognition, gait identification and soft-biometric features. Subsequently, work was done on the fusion of these models to create a robust and accurate system. Finally, a validation, adjustment and documentation phase was carried out. The proposed system proved to be efficient in preliminary tests, presenting results in accordance with the literature review of the project.

Keywords: person Re-Identification, biometric features, soft-biometric features, computer vision, machine learning, scrum.

Capítulo I

Introducción

Propósito y contextualización del tema

Justificación del interés de la investigación

Actualmente, la Visión por Computador moderna se especializa en los sistemas de Video Vigilancia, debido al incremento de la demanda de seguridad pública y la propagación de cámaras de seguridad en multitud de espacios públicos y privados, siendo la seguridad una de las prioridades de la sociedad moderna, porque benefician el bienestar de las personas, como sus hogares, casas, negocios, edificios, centros, escuelas, centros de salud, público y lugares privados.

El aumento de las ciudades inteligentes ha incrementado la demanda de sistemas de Video Vigilancia inteligentes. Debido a que estos sistemas se basan en la Re-Identificación de personas, esta tecnología ha tomado relevancia. La Re-Identificación de personas emplea el análisis de características biométrías y soft-biométrías de los individuos capturados por las cámaras (*How Face Recognition Technology and Person Re-Identification Technology Can Help Make Our World Safer and Secure | NEC*, n.d.; Hua et al., 2022). Entre las características biométrías, la forma de andar ha mostrado tener especial relevancia. Varios estudios han demostrado que los humanos pueden reconocer a las personas sólo por su forma de andar, siempre que se disponga de mediciones adecuadas y de buena calidad de los patrones de movimiento observados (Chugchilan Hinojosa & Guanochanga Farfán, 2023; Cuzzolin, 2010; Mandar et al., 2021; Nambiar et al., 2019). Para aplicaciones de vigilancia, la forma de andar también resulta atractiva, porque no requiere la colaboración activa de las personas y es difícil de simular o falsear.

Este trabajo integra la forma de caminar, al ser captada por distintas cámaras, como una nueva característica biométrica la cual podría aumentar el porcentaje de Re-Identificación de personas, ya que tiene una fuerte correlación con la identidad de la persona. Así también, cabe mencionar que el

presente trabajo es la continuación de otros trabajos de Re-Identificación de personas realizados en la ESPE sede Latacunga. Este sistema se basa en diversas características biométricas y softbiométricas para complementar su funcionamiento. Las características biométricas utilizadas para este sistema son el: rostro y la forma de caminar, mientras que las características softbiométricas se encuentran la silueta corporal, la textura y el color de la vestimenta. El análisis se lleva a cabo en un entorno controlado, utilizando técnicas de Visión por Computador y modelos y/o algoritmos de Machine Learning para procesar los datos recopilados por las cámaras.

Con el fin de validar este enfoque se realizará comparaciones entre los resultados obtenidos utilizando características biométricas y soft-biométricas por separado, así también, evaluar cómo la combinación de ambos tipos de características puede incrementar la precisión en el reconocimiento de una persona.

Para la implementación de este sistema se utilizarán técnicas avanzadas de visión por computadora, tales como redes neuronales, modelos holísticos y no holísticos, descriptores, clasificadores y técnicas de fusión de modelos de Aprendizaje Profundo. La elección de estas técnicas parte de trabajos previamente realizados referente a Re-Identificación de personas en la ESPE Sede Latacunga, así también, mediante la revisión de literatura especializada en el tema se pudo identificar que este conjunto de técnicas son las más utilizadas y las que mejores resultados han obtenido.

Para ello, se utilizará un conjunto de tecnologías basadas en el lenguaje de programación Python, mismo que es especializado para temas relacionados con la Visión por Computadora e Inteligencia Artificial, debido a que posee un extenso conjunto de librerías y marcos de trabajo, así como una vasta comunidad de investigadores y colaboradores alrededor del mundo.

Objetivos Generales y Específicos

A continuación, se plantean los objetivos generales y específicos del presente trabajo:

Objetivo General

- Desarrollar un sistema de Re-Identificación de personas a través de las características biométricas: facial y la forma de caminar y softbiométricas: silueta corporal, textura y color de la vestimenta, con el uso de modelos no holísticos, aplicando técnicas computacionales avanzadas.

Objetivos Específicos

- Conocer y estudiar sobre técnicas computacionales avanzadas utilizadas en la Re-Identificación de personas utilizando características biométricas: facial y la forma de caminar y softbiométricas: silueta corporal, textura y color de la vestimenta.
- Implementar un sistema de Re-Identificación de personas a través de las características biométricas: facial y la forma de caminar, y características softbiométricas: silueta corporal, textura y color de la vestimenta para personas, utilizando técnicas computacionales avanzadas.
- Validar los resultados, analizar errores y ajustar los modelos del sistema de videovigilancia.

Metodología

El propósito principal de este estudio consiste en desarrollar un sistema de Re-Identificación de personas mediante el uso de las características biométricas, como el rostro y la forma de caminar, y características softbiométricas, tales como la silueta corporal, textura y color de la vestimenta. Se emplearán modelos no holísticos, aplicando técnicas computacionales avanzadas para lograr este objetivo. En primer lugar, se comienza examinando el tema de la Re-Identificación de personas mediante el uso de características biométricas (rostro y forma de caminar) y softbiométricas (silueta corporal, textura y color de la vestimenta), a través de la búsqueda en la literatura científica. Esta investigación se realizará utilizando un enfoque sistemático para establecer una base teórica inicial.

Esto es seguido por una investigación sobre la Re-Identificación de personas aplicando modelos holísticos y no holísticos, técnicas de Visión por Computadora, y Machine y modelos y/o algoritmos de

Machine y/o Deep Learning. La investigación se realizará empleando la búsqueda y revisión de publicaciones en las bases de datos Scopus e IEEExplore.

Posteriormente, se analizará datasets que incluyan imágenes o videos de personas con variaciones en poses, iluminación, ángulos de cámara, y vestimenta. Esto permitirá determinar las condiciones del entorno sobre el cual se ejecutarán las pruebas, para mejores resultados. A continuación, se generan un conjunto de datos con los videos capturados por el circuito cerrado de cámaras no superpuestas utilizadas en el proyecto. De estos conjuntos de datos se extraen los rasgos faciales, la textura y el color de la ropa, la forma del cuerpo y las características de la marcha.

Estas características son utilizadas para generar descriptores, implementar modelos y algoritmos de Machine Learning que permitan la identificación de personas en diferentes ubicaciones (predicciones). Posteriormente, se aplican métricas de precisión para evaluarlos, utilizando los conjuntos de datos previamente creados y/o datos de prueba (Videos). Además, se realizarán variaciones en los hiperparámetros de los modelos seleccionados, los cuales podrían mejorar el proceso de aprendizaje, con el objetivo de comparar y analizar las predicciones obtenidas.

Se utilizara el descriptor Haar Cascade porque es un método comúnmente utilizado para identificar rostros. El método Haar Cascade utiliza características locales, como bordes y esquina, para encontrar regiones faciales en una imagen. Una vez identificadas estas regiones faciales, algoritmos como Local Binary Patterns (LBP) o Histogram of Oriented Gradients (HOG) son utilizados para obtener rasgos faciales distintivos.

Utilizamos el descriptor en cascada de Haar porque es una técnica muy utilizada para el reconocimiento facial (Viola & Jones, 2001).. El enfoque en cascada de Haar utiliza características locales como bordes y esquinas para localizar regiones de caras en una imagen. Una vez identificadas estas regiones faciales, se utilizan técnicas de extracción de características faciales como los gradientes

orientados por histograma (HOG) (Ahonen *et al.*, 2006) y los patrones binarios locales (LBP) (Dalal & Triggs, 2005a).

Para el análisis de la marcha el descriptor GEI (Gait Energy Image) resulta ampliamente utilizado (Goffredo, 2014). Este descriptor consiste en capturar las secuencias de imágenes mientras una persona camina. A continuación, se promedian estas imágenes para crear una representación de la forma de andar. Esta representación captura características relevantes de la forma de caminar de una persona, como la longitud de paso, el ángulo de la rodilla y el ritmo (Han & Bhanu, 2006).

La combinación de las características faciales y la forma de caminar se realiza empleando el clasificador SVM (Support Vector Machine). Este clasificador es una técnica de aprendizaje supervisado ampliamente utilizada en tareas de clasificación de Videos (Perea-Ortega *et al.*, 2009). Este se entrena utilizando un conjunto de datos previamente etiquetados que incluye en este caso características faciales y de la marcha de personas conocidas (Cortes & Vapnik, 1995a). Durante el entrenamiento, el clasificador SVM busca un hiperplano perfecto para dividir varias clases de personas en el espacio de características.

Para capturar la silueta corporal, se aplica el enfoque de Histogram of Oriented Gradients (HOG) (Dalal & Triggs, 2005a). Este método calcula y representa las distribuciones de gradientes en una imagen para obtener información sobre la forma y la estructura del cuerpo. El HOG ha demostrado ser efectivo en la detección y descripción de formas humanas en imágenes.

En cuanto al reconocimiento del color y la textura de la vestimenta, se utiliza el enfoque de detección y clasificación de objetos YOLO (You Only Look Once) (Redmon *et al.*, 2016a). YOLO es un modelo de detección de objetos en tiempo real que identifica y localiza diferentes clases de objetos en una imagen. Para el reconocimiento del color, se utiliza la transformación a la representación de espacio de color HSV (Hue, Saturation, Value), que proporciona información sobre la tonalidad y la intensidad del color de la vestimenta (Sharma, 2019). Mediante el descriptor Local Binary Patterns (LBP), se aborda

la textura de la ropa (Ojala *et al.*, 2002a), con este se captura patrones locales y variaciones en la textura de la ropa.

Para combinar las tres características softbiométricas, utilizamos un enfoque de fusión de características. En este caso, el clasificador SVM (Support Vector Machine) para fusionar las características de silueta corporal, color y textura de la vestimenta (Cortes & Vapnik, 1995a). Para entrenar el clasificador SVM se utiliza un conjunto de datos etiquetados con características extraídas de personas conocidas. Durante el entrenamiento, el clasificador SVM busca encontrar un hiperplano óptimo que separe las diferentes clases de personas en el espacio de características combinadas.

Una vez entrenado el clasificador SVM, es utilizado para predecir la identidad de una persona desconocida utilizando las características softbiométricas combinadas. El modelo SVM compara las características de la persona desconocida con las características de las personas conocidas en el conjunto de entrenamiento y asigna una etiqueta o clase basada en la similitud de las características.

La combinación de los enfoques biométricos y soft-biométricos se realiza en primer lugar incorporando el modelo de reconocimiento del rostro con los modelos soft-biométricos combinados. A continuación, se implementará la forma de caminar, modelo que posteriormente será fusionado con los combinados previamente.

La validación de resultados se lleva a cabo mediante pruebas empleando el conjunto de datos de prueba. Estas pruebas se enfocan en alcanzar el nivel de precisión que se encuentre dentro del rango de precisión encontrada en la literatura.

En la revisión de literatura científica enfocada a Re-Identificación de personas las características biométricas como el rostro han alcanzado altos porcentajes de reconocimiento. Por ejemplo, en un estudio de reconocimiento facial utilizando técnicas basadas en deep learning, se logró una tasa de reconocimiento del 99.63% en un conjunto de datos desafiante (*How Face Recognition Technology and Person Re-Identification Technology Can Help Make Our World Safer and Secure* | NEC, n.d.; Sun et al.,

2014a). En otro estudio sobre reconocimiento de la forma de caminar, se informó una tasa de reconocimiento del 95.7% utilizando un enfoque basado en análisis de características 3D (Das *et al.*, 2014).

Por otro lado, en términos de Re-Identificación de personas, las características soft-biométricas han mostrado resultados alentadores. En una investigación que utilizó características de silueta corporal, se logró una tasa de reconocimiento del 92.3% en un conjunto de datos de personas en movimiento (Gafurov, 2011). En otro estudio el reconocimiento basado en características de color y textura de la ropa presentó una tasa de reconocimiento del 85.4% utilizando un enfoque de Machine Learning.

L Los porcentajes de reconocimiento son propensos a variar de acuerdo con el conjunto de datos utilizado, técnicas y condiciones de captura de las imágenes empleadas en la Re-ID. Los resultados de la investigación revisada indican que la técnica de combinación propuesta de características biométricas y soft-biométricas podría alcanzar altas tasas de reconocimiento. Esto se debe a que en este trabajo se usa un circuito cerrado de cámaras las cuales no están solapadas, que crea un ambiente controlado sin cambios significativos en el entorno como cambios de luz u oclusiones.

Capítulo II

Marco Teórico

Introducción al Marco Teórico

En este capítulo se describen las técnicas y algoritmos utilizados para el desarrollo del proyecto de Re-Identificación de personas. La Re-Identificación de personas en los campos de la Visión por Computadora y el Machine Learning implica localizar y seguir a una persona concreta que aparece en una red de cámaras conectadas, incluso cuando la apariencia de esa persona experimenta cambios importantes. En esencia, implica asociar correctamente las imágenes o videos de una persona capturada en diferentes cámaras y momentos.

Una estrategia comúnmente utilizada para la Re-Identificación de personas es extraer características discriminativas de las imágenes o videos, que capturan las características distintivas de una persona. Estas características pueden incluir aspectos como la apariencia física, el color de la ropa, la textura, la forma y la postura. Luego, se utilizan algoritmos de comparación para calcular la similitud entre las características extraídas de diferentes imágenes o videos, con el fin de asociar correctamente a una persona en diferentes cámaras.

El uso de características biométricas y soft-biométricas para extraer patrones que permitan el reconocimiento de una persona a través de sus características particulares es uno de los métodos más populares para la Re-Identificación de personas. Estos enfoques han demostrado mejoras significativas en la precisión y robustez de la Re-Identificación de personas (Zheng *et al.*, 2016).

Características Biométricas y Soft-biométricas

La Re-Identificación de personas implica el uso de varias estrategias para localizar y detectar a una persona a través de múltiples cámaras, donde la combinación de datos biométricos y soft-biométricos se emplea frecuentemente para mejorar la precisión del reconocimiento. Las características

biométricas: reconocimiento facial y la forma de andar, así como las características soft-biométricas: silueta corporal, la textura y el color de la ropa, se analizan en los siguientes apartados.

Características Biométricas

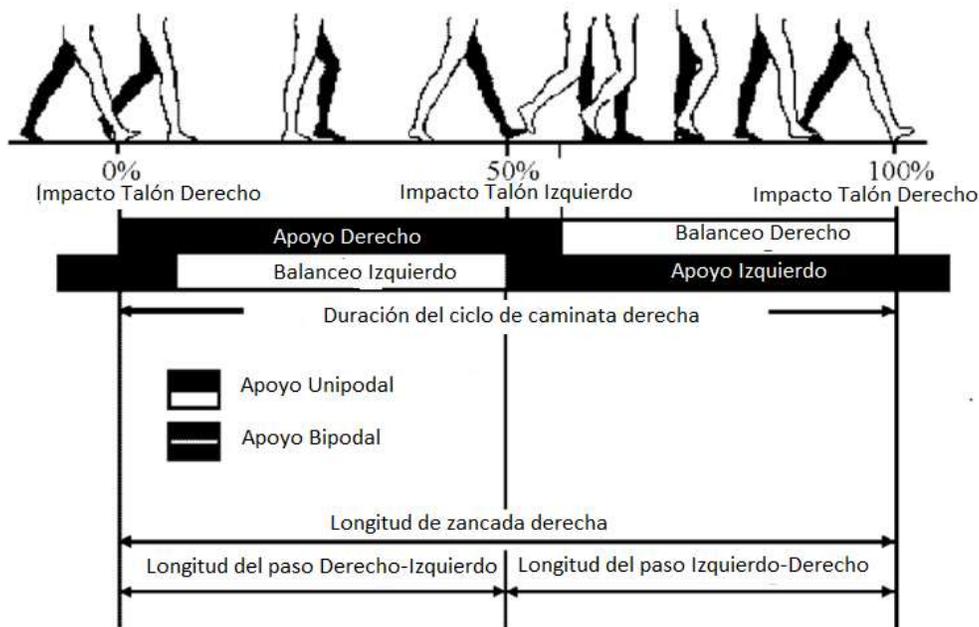
Se refieren a aquellos rasgos físicos que son únicos para cada individuo, estos rasgos pueden ser medidos y utilizados para identificar a una determinada persona. Ejemplos comunes de características biométricas incluyen huellas dactilares, patrones de retina o iris, geometría de la mano, y características faciales. En el contexto de esta investigación, se han destacado el rostro y la forma de caminar como características biométricas clave (Jain *et al.*, 2004).

El reconocimiento por la forma del rostro consiste en la determinación de características únicas como la longitud de separación entre cada ojo, las dimensiones de la nariz, la posición de los pómulos, entre otros. Estos rasgos son difíciles de alterar y ofrecen una forma precisa de identificación (W. Zhao *et al.*, 2003).

Por otra parte, la forma en que una persona camina, conocida como "gait", es única y puede ser utilizada como una característica biométrica. Esta técnica se basa en el movimiento y la forma del cuerpo al caminar (ver Figura 1), y aunque puede ser afectada por factores externos como la ropa o el calzado, sigue siendo un indicador valioso para la identificación puesto que es único de cada individuo (Niyogi & Adelson, 1994).

Figura 1

Ciclo de la caminata



Nota. El gráfico representa el ciclo de la marcha donde se aprecia las longitudes de paso y zancada durante un ciclo. Tomado de *Motion Analysis and Tracking* (p.231), por M.S. Nixon, J.N. Carter, D. Cunado, P.S. Huang, S.V. Stevenage, 2006, University of Southampton, Southampton, UK.

Los rasgos soft-biométricos, que no son específicos de cada individuo, pero resultan especialmente útiles en circunstancias en las que los rasgos biométricos no completamente evidentes, son utilizados en esta investigación para aumentar la robustez del sistema de Re-Identificación de personas.

Características Soft-biométricas

Las características soft-biométricas como el color y la textura de la vestimenta, la silueta corporal, el género, la altura, y el color del cabello, entre otros (Dantcheva *et al.*, 2011), a diferencia las biométricos tradicionales, no siempre son distintivas de una persona, pero pueden ayudar al proceso de identificación aportando más detalles. Estas características suelen ser más fáciles de obtener y menos invasivas.

La Figura 2 representa un claro ejemplo donde las características biométricas no son distinguibles fácilmente, no obstante, las características Soft-biométricas brindan una oportunidad para identificar a un determinado individuo.

Figura 2

Características soft-biométricas



Nota. Las características Soft-biométricas permiten capturar patrones que mejoran la Re-Identificación de un individuo en diferentes escenas, son especialmente útiles cuando no se pueden obtener características biométricas debido a las variaciones del entorno (iluminación, cambio de pose, oclusión).

Tomado de *Computer Vision – ACCV 2020* (p. 71), por Xuelin Qian, Wenxuan Wang, Li Zhang, Fangrui Zhu, Yanwei Fu, Tao Xiang, Yu-Gang Jiang, and Xiangyang Xue, 2021, Springer International Publishing.

La silueta se refiere al contorno del cuerpo humano. Aunque no es única para cada individuo, puede proporcionar información valiosa, especialmente cuando se combina con otras características (Dantcheva *et al.*, 2011).

Así también, la ropa que una persona lleva, incluyendo su color y textura, puede ser utilizada para ayudar en la identificación, especialmente en entornos de vigilancia donde otras características pueden no ser claramente visibles.

Importancia en la Re-Identificación

La Re-Identificación de personas se pueden complementar a través de la integración de características biométricas y soft-biométricas, lo que proporciona un método sólido y versátil (Hua et al., 2022). Las características soft-biométricas puede ser especialmente útiles en circunstancias en las que los rasgos biométricos no son fácilmente evidentes o son difíciles de obtener, aunque los rasgos biométricos ofrecen un alto nivel de precisión debido a su naturaleza distintiva no siempre son fáciles de adquirir. Por ejemplo, en un entorno de vigilancia, la vestimenta de una persona o su silueta pueden ser más evidentes que sus rasgos faciales. Al fusionar la información obtenida de ambas características, se puede mejorar significativamente la precisión y la eficiencia del proceso de Re-Identificación (Ming et al., 2022b).

Ahora, la Re-Identificación de personas hace uso de técnicas para capturar patrones de una imagen, estas técnicas consisten en tomar el objeto de interés como un todo y procesarlo o tratar a la imagen mediante la segmentación de zonas que representen ciertas características, cuando se procesa a una imagen como un todo estamos hablando de una técnica holística, mientras que cuando se segmentan las zonas de interés de una imagen estamos refiriéndonos a una técnica no holística, el sistema de Re-Identificación de personas hace uso de estas dos técnicas.

Técnicas Holísticas y No Holísticas

En el ámbito de la visión por computadora, especialmente en tareas relacionadas con la identificación y Re-Identificación de personas, se utilizan diferentes enfoques para procesar y analizar imágenes. Estos enfoques se pueden categorizar en modelos holísticos y no holísticos. La elección entre uno y otro depende en gran medida de la naturaleza de la tarea y de los datos disponibles (Ye et al., 2013).

La elección entre enfoques holísticos y no holísticos en visión por computadora a menudo se basa en la naturaleza de la tarea y los datos disponibles. Ambos enfoques tienen sus propias ventajas y desventajas, lo que los hace más adecuados para ciertas aplicaciones.

Modelos Holísticos

Estos modelos consideran la imagen en su totalidad, sin descomponerla en partes o características individuales. Son especialmente útiles cuando se necesita una visión general o cuando las características individuales no aportan significativamente al resultado final.

Los modelos holísticos tienen un procesamiento más rápido debido a que requieren un menor preprocesamiento de la imagen, no obstante, son de menor precisión en tareas que requieren un análisis detallado como en la Re-Identificación de personas, así como son muy susceptibles a los cambios o variaciones en el entorno como: iluminación, postura u oclusión.

Para el reconocimiento facial se emplean modelos holísticos, es decir, se hace uso de las imágenes que contenga la región del rostro (ver Figura 3), sin descomponerlo en características individuales (Turk & Pentland, 1991), mientras que para el reconocimiento de una persona mediante su forma de caminar se hace uso de modelos no holísticos como se detallará holísticos (ver sección: Modelos No Holísticos).

Figura 3

Representación de modelo holístico



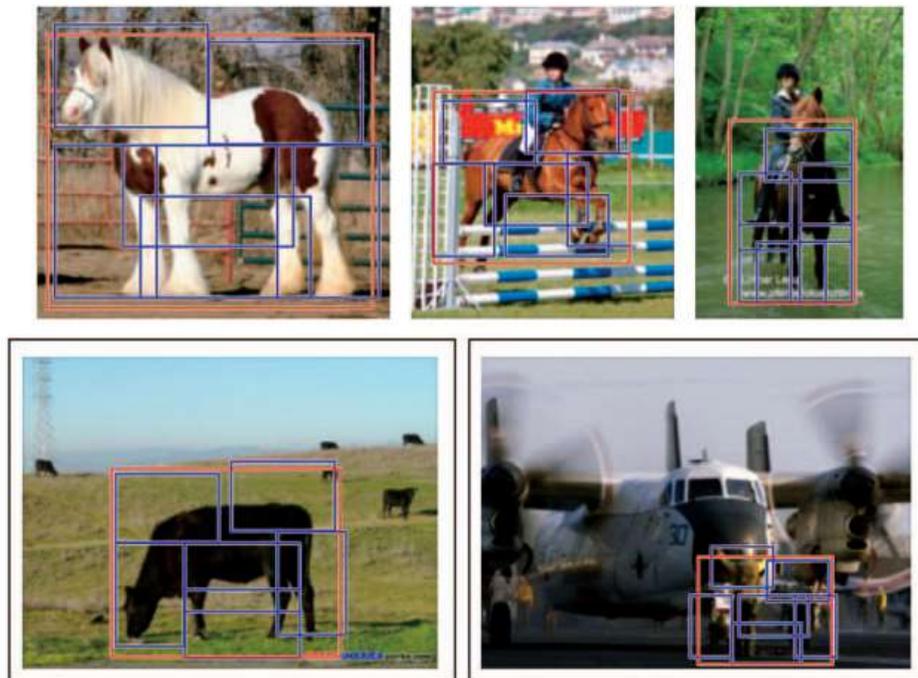
Nota. Representación de un modelo holístico que toma la imagen completa del rostro de una persona, sin descomponerlo ni dividirlo en secciones. Tomado de *Face Recognition Using Eigenfaces* (p. 588), por Matthew A. Turk, Alex P. Pentland, 1991, Massachusetts Institute of Technology.

Modelos No Holísticos

Los modelos no holísticos descomponen la imagen en características o partes específicas (ver Figura 4). Un ejemplo prominente de este enfoque es el modelo DPM (Deformable Part Models). DPM es una técnica que busca detectar y modelar partes deformables de un objeto. Utiliza un conjunto de partes y establece relaciones espaciales entre ellas. Cada parte se representa mediante un filtro y se busca en la imagen a diferentes escalas y posiciones (Felzenszwalb *et al.*, 2010).

Figura 4

Representación de modelo no holístico



Nota. Representación de un Modelo No Holístico que identifica caballos, donde en los recuadros en rojo el Modelo No Holístico a reconocido un caballo y los recuadros en azul son las regiones de interés que

analiza el Modelo No holístico para determinar si se ha detectado un caballo. Tomado de *Object Detection with Discriminatively Trained Part-Based Models* (p. 7), por David Forsyth, 2014, University of Illinois.

DPM utiliza un algoritmo basado en la búsqueda de máxima puntuación entre las partes y la imagen, considerando las deformaciones. El objetivo es encontrar la configuración de partes que maximice la correspondencia con la imagen.

En investigaciones previas, DPM ha demostrado ser altamente eficiente en la detección de objetos, alcanzando precisiones superiores al 80% en datasets estándar como PASCAL VOC (Felzenszwalb *et al.*, 2010). Así pues, los modelos no holísticos tienen mayor precisión al centrarse en detalles específicos y son especialmente útiles en situaciones donde exista variación en postura, iluminación u oclusión.

No obstante, los modelos no holísticos requieren un mayor preprocesamiento, lo cual resulta en modelos que son computacionalmente costosos. En el presente trabajo se hace uso de esta técnica no holística para detectar el torso de una persona para extraer los contornos y así obtener la silueta corporal, lo cual es esencial para los modelos de que utilizan características Soft-biométricas que se analizarán posteriormente.

Ahora una vez definidas las técnicas de procesar las imágenes y el conjunto de características Biométricas y Soft-biométricas a utilizar es necesario detallar las Técnicas de Visión por Computadora mediante las cuales se pueden extraer patrones que permiten identificar a una persona.

Técnicas de Visión por Computadora

El objetivo de Visión por Computador, una rama de la inteligencia artificial y el procesamiento de imágenes, es dotar a las máquinas la capacidad de comprender e interpretar los datos visuales del entorno de forma análoga a la percepción humana (Marr, 2010). Su objetivo fundamental es imitar la forma en que las personas ven y comprenden el mundo que les rodea.

El flujo de trabajo en la visión por computadora para lograr el reconocimiento de una persona implica varias etapas. En primer lugar, se lleva a cabo la adquisición de imágenes o videos que contienen la información visual de interés. Estas imágenes pueden ser capturadas mediante cámaras o provienen de fuentes de video. A continuación, se procede con el procesamiento de imágenes, donde se aplican técnicas y algoritmos para mejorar la calidad de las imágenes y eliminar el ruido. Esto puede incluir operaciones como el filtrado para suavizar la imagen y reducir el ruido, la corrección de la iluminación y el contraste, y la normalización para asegurar una representación consistente de las imágenes.

Después del procesamiento de imágenes, se realiza la detección de la persona en las imágenes. Para esto, se utilizan algoritmos de detección de objetos, como el algoritmo de Viola-Jones (Viola & Jones, 2001) y los clasificadores basados en Haar cascade (Viola & Jones, 2004). Para identificar una persona que se encuentra en una imagen estos algoritmos dependen del reconocimiento de determinadas características, como bordes y patrones visuales para identificarlo.

Una vez que se ha detectado la persona, se procede a la extracción de características relevantes para el reconocimiento. Esto implica la identificación y extracción de atributos distintivos de la persona, como el rostro y la forma de caminar. Para la identificación facial se utilizan técnicas como el análisis de rasgos faciales y la coincidencia de patrones, incluidos los Local Binary Patterns (LBP) (Ahonen *et al.*, 2006). Para la forma de caminar, se emplean técnicas como el uso de redes neuronales convolucionales (CNN) (Shi *et al.*, 2023).

Una vez extraídas las características, se pasa a la etapa de clasificación, en la que se emplea una técnica de clasificación como las máquinas de vectores soporte (SVM) (Cortes & Vapnik, 1995b), para asignar las características a una clase específica, como la identidad de la persona o la categoría de interés.

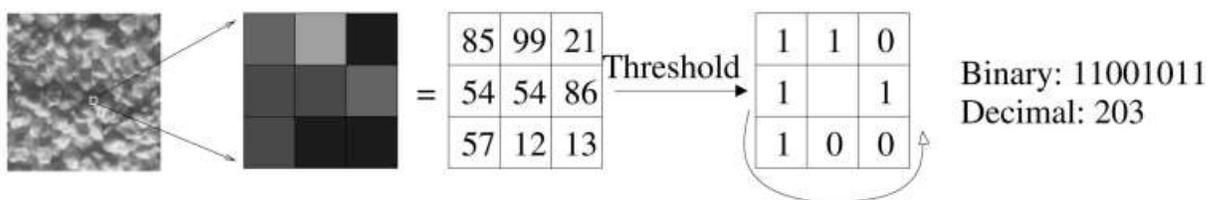
Local Binary Pattern (LBP)

Una estrategia eficaz para la descripción de texturas es el método de los patrones binarios locales (LBP), que se basa en la noción de que las texturas locales de una imagen pueden caracterizarse por la distribución de patrones diminutos, como bordes, puntos y esquinas, entre otros (Ojala *et al.*, 1996).

La técnica LBP compara el valor de un píxel central con sus vecinos en una pequeña región alrededor de él (ver Figura 5). Si el valor del píxel vecino es mayor que el del píxel central, se le asigna un valor de 1, y si es menor, se le asigna un valor de 0. De esta manera, para cada píxel central, se obtiene un patrón binario local.

Figura 5

Asignación de valores a píxeles mediante LBP



Nota. Representa la asignación de valores binarios a una determinada región de una imagen mediante LBP. Tomado de *Face Description with Local Binary Patterns: Application to Face Recognition* (p. 2038), por Timo Ahonen, 2006, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (12).

Una vez que se ha calculado el patrón binario local para cada píxel, se genera un histograma que cuenta la frecuencia de ocurrencia de cada patrón. Este histograma se puede utilizar como una representación compacta de la distribución espacial de los patrones locales en la imagen.

El LBP ha demostrado ser eficaz en una variedad de aplicaciones en visión por computadora, especialmente en la clasificación de texturas. Dado que es invariante a cambios monotónicos en la iluminación y es computacionalmente eficiente, el LBP es adecuado para aplicaciones en tiempo real.

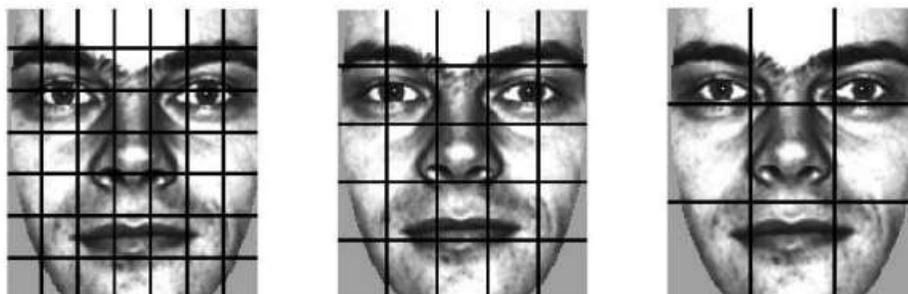
Además, en el contexto de la Re-Identificación, el LBP puede ser utilizado para capturar detalles finos en imágenes, como la textura de la ropa o características faciales, lo que puede ser crucial para distinguir entre individuos.

Existen variaciones del algoritmo LBP, como LBPH (Local Binary Patterns Histograms), que extienden el enfoque original permitiendo el uso de vecindades más grandes (ver Figura 6) y considerando una estrategia de interpolar los valores de los píxeles en la vecindad. Esto mejora la robustez del descriptor y su capacidad para capturar detalles más finos en las texturas o características faciales, entonces al igual que LBP la técnica LBPH se utiliza para construir un histograma de patrones binarios locales en una imagen facial. Este histograma luego se utiliza como una representación compacta de la apariencia facial, que se puede comparar con otros rostros utilizando técnicas de clasificación o reconocimiento de patrones.

En el reconocimiento facial, el descriptor LBP ha demostrado ser efectivo en la extracción de características locales distintivas, como patrones de textura, que son útiles para la identificación de personas.

Figura 6

Selección de regiones de una imagen facial



Nota. Representa la división de regiones para la captura de detalles en una imagen facial mediante LBP, a más divisiones permite obtener detalles más locales de la imagen, mientras que con menos divisiones se capturan patrones globales de una imagen. Esto implica que las imágenes con mayores celdas son

más robustas antes cambios en la iluminación u orientación de la cara. Tomado de *Face Description with Local Binary Patterns: Application to Face Recognition* (p. 2038), por Timo Ahonen, 2006, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28 (12).

La implementación de la técnica LBP en el reconocimiento de rostros conlleva la elección de una vecindad, esquema de codificación binaria y el cálculo del histograma. Además, es importante considerar el formato de la imagen y el número de píxeles a utilizar para garantizar una representación efectiva de las características locales.

Al examinar los patrones de intensidad de los píxeles, el algoritmo LBP puede captar aspectos locales y texturales de una imagen. La información de los patrones locales se captura en forma de histograma, que representa la distribución de los diferentes patrones en la imagen.

Por otra parte, en esta investigación se usa el algoritmo de Viola-Jones para detectar regiones de interés del rostro de una persona, este algoritmo es muy utilizado porque proporciona tasas altas de reconocimiento y es eficiente computacionalmente.

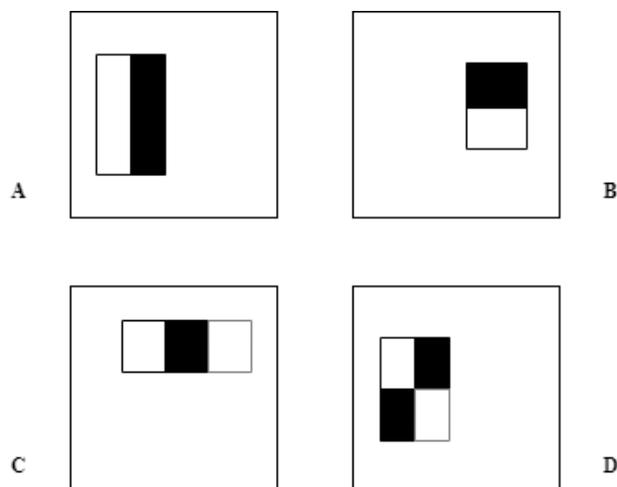
Algoritmo de Viola-Jones

En el campo de la Visión por Computadora, el algoritmo Viola-Jones, desarrollado por Viola y Jones en 2001, es una técnica muy apreciada para la detección de objetos, incluido el reconocimiento facial y la detección de personas (Viola & Jones, 2001).

El algoritmo se basa en el análisis de características de imágenes conocidas como características de Haar. Estas características son patrones rectangulares de cambios de intensidad de píxeles en una imagen como se muestra en la Figura 7. El algoritmo utiliza una técnica de Machine Learning llamada clasificador Adaboost para combinar de manera eficiente un gran número de características y construir un detector robusto (Viola & Jones, 2001).

Figura 7

Características de Haar



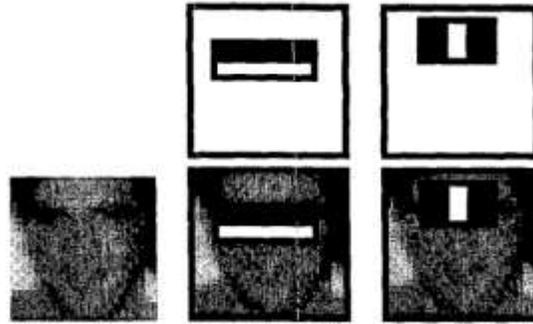
Nota. Representan rectángulos de características que son usados en la detección de rostros, donde A y B representan rectángulos de 2 características de Haar, mientras que C representa rectángulos con 3 características de Haar y D representa rectángulos con 4 características de Haar. Tomado de *Rapid object detection using a boosted cascade of simple features* (p. 1-512), por Paul Viola, 2001, IEEE Computer Society.

El proceso del algoritmo comienza con la conversión de la imagen a escala de grises, ya que las características de Haar se basan en cambios de intensidad. Luego, se desliza una ventana rectangular sobre la imagen a diferentes escalas y se calculan las respuestas de las características de Haar en cada posición. Estas respuestas se utilizan para determinar la presencia o ausencia de un rostro en la región evaluada (Viola & Jones, 2001).

El algoritmo de Viola-Jones se utiliza por etapas, durante el entrenamiento se recogen que contienen rostro e imágenes que no la contengan. Las características de Haar se calculan en estas imágenes y se utiliza el clasificador Adaboost para seleccionar las características más discriminativas tal y como se muestra en la Figura 8. Luego, se construye una cascada de clasificadores que se aplicará en la etapa de detección (Viola & Jones, 2001).

Figura 8

Regiones identificadas por el clasificador Adaboost



Nota. Representa las regiones que se reconocen como rostros de personas debido a la intensidad de los píxeles en esas determinadas áreas utilizando el algoritmo Viola-Jones. Tomado de *Rapid object detection using a boosted cascade of simple features* (p. 1-514), por Paul Viola, 2001, IEEE Computer Society.

Durante la etapa de detección, la cascada de clasificadores se aplica a la imagen de entrada utilizando una estrategia en cascada. Esto significa que se utilizan clasificadores más complejos y computacionalmente costosos solo en las regiones donde se ha detectado la presencia de un objeto potencial. Esto permite un procesamiento más rápido y la reducción del número de falsos positivos (Viola & Jones, 2001).

Así también, el algoritmo Viola-Jones presenta varias ventajas en el reconocimiento facial y la detección de personas. Es rápido y eficiente, lo que lo hace adecuado para aplicaciones en tiempo real. Además, es robusto ante variaciones en la apariencia y puede detectar rostros en diferentes poses, escalas e iluminaciones. Sin embargo, el algoritmo puede ser sensible a cambios drásticos en la iluminación y puede generar falsos positivos en escenarios con objetos similares a los rostros (Viola & Jones, 2001).

Así pues, una parte fundamental que compone el flujo de la Re-Identificación de personas es la cuando el modelo implementado requiere clasificar entre determinadas clases, es decir, decidir si se detectó a la persona A o persona B, para ello las máquinas de vectores de soporte son ampliamente

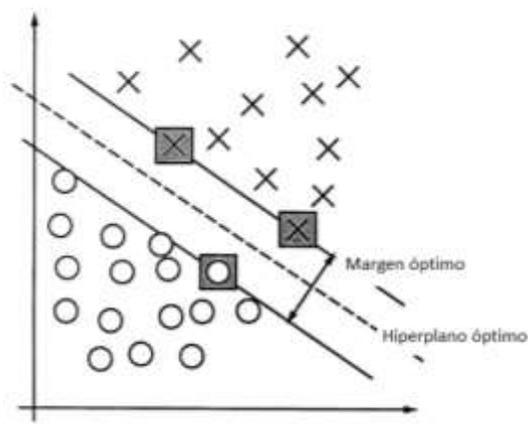
utilizadas porque han resultado eficientes y computacionalmente viables para este tipo de tareas como se explicara a continuación.

Support Vector Machine (SVM)

El clasificador de máquinas de vectores de soporte (Support Vector Machine), es un algoritmo de Machine Learning utilizado en visión por computadora. Su objetivo principal es encontrar un hiperplano que mejor divida un conjunto de datos en clases. Cuando se trata de datos que puedan separarse linealmente, se identifica el hiperplano que tenga la mayor distancia (margen) entre los puntos de datos más cercanos de cada clase, esta técnica se conoce como vector de soporte (Cortes & Vapnik, 1995b) tal y como se muestra en la Figura 9.

Figura 9

Representación de SVM



Nota. Representa una clasificación SVM en un espacio bidimensional (2 posibles clasificaciones), en esta figura se aprecia el concepto de SVM en el hiperplano y el margen de división entre 2 clases plenamente identificadas. Tomado de *Support-vector networks* (p. 275), por Cortes & Vapnik, 1995, *Machine Learning*, 20 (3).

En el contexto del reconocimiento facial y la detección de personas, el clasificador SVM se utiliza para separar y distinguir entre diferentes clases de objetos, como rostros o personas. Para ello, se requiere un conjunto de datos de entrenamiento que contengan ejemplos positivos y negativos de la

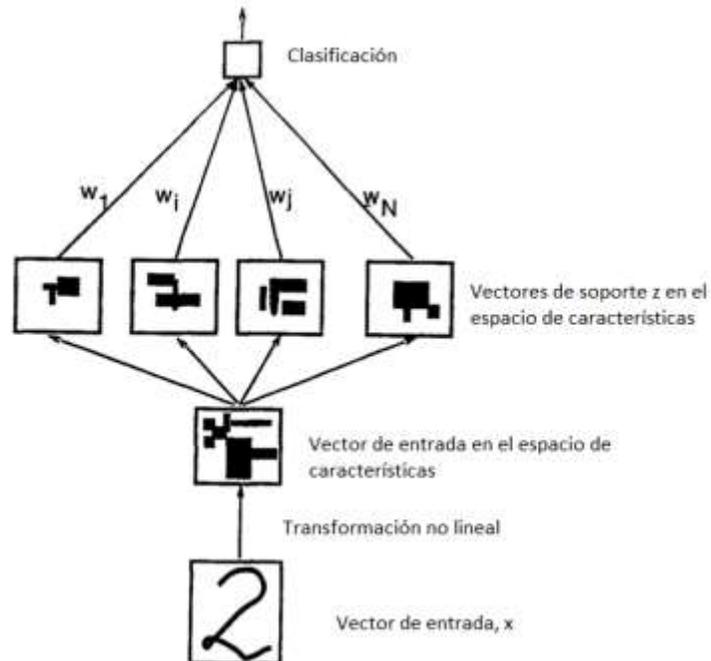
clase objetivo. Este conjunto de datos es utilizado para entrenar el clasificador SVM y establecer una función de decisión que pueda separar eficientemente las clases (Osuna *et al.*, 1997).

Mediante una función de núcleo, los datos de entrada se transforman en un espacio de mayor dimensión como parte del proceso de clasificación SVM. Esta transformación permite encontrar un hiperplano en el nuevo espacio que maximice el margen entre las clases, lo que se traduce en una mejor capacidad de generalización y robustez en la clasificación (Cortes & Vapnik, 1995b), este proceso de clasificación se evidencia en la Figura 10.

El algoritmo SVM consta de varias etapas. En primer lugar, se realiza una selección y preprocesamiento de características relevantes de las imágenes, como el reconocimiento facial o los descriptores de características de personas. A continuación, se utiliza un conjunto de entrenamiento etiquetado para entrenar el clasificador SVM y se modifican los parámetros del modelo para producir la mejor separación de clases posible. Finalmente, se utiliza el modelo entrenado para clasificar nuevos datos de entrada (Osuna *et al.*, 1997).

Figura 10

Clasificación en una red de vectores de soporte (SVM)



Nota. Representación del proceso de clasificación en una red de vectores de soporte (SVM), en la cual se separan las características más representativas en un hiperplano dado un vector de entrada. Tomado de *Support-vector networks* (p. 277), por Cortes & Vapnik, 1995, *Machine Learning*, 20 (3).

El clasificador SVM presenta varias ventajas en el reconocimiento facial y la detección de personas. Es capaz de manejar conjuntos de datos complejos y de alta dimensión, lo que lo hace adecuado para la clasificación de características visuales. Además, el SVM es eficiente en términos de tiempo de entrenamiento y puede proporcionar una buena generalización en diferentes escenarios (Cortes & Vapnik, 1995b).

Sin embargo, el SVM también tiene algunas limitaciones. Requiere una selección adecuada del kernel y ajuste de parámetros para obtener resultados óptimos, lo cual puede ser un proceso no trivial. Además, su rendimiento puede verse afectado por desequilibrios en los conjuntos de datos o por la presencia de ruido o datos atípicos (Cortes & Vapnik, 1995b).

El clasificador SVM es parte de las técnicas de Machine Learning debido a su capacidad para aprender patrones complejos y realizar clasificaciones precisas en datos de entrada. A través del

proceso de entrenamiento y ajuste de parámetros, el SVM logra construir una función de decisión que separa las clases objetivo de manera óptima (Osuna *et al.*, 1997).

Redes Neuronales Convolucionales (CNN)

Las redes neuronales convolucionales (CNN) es un método innovador en el campo de Visión por Computador. Estas redes se han destacado por su eficacia en tareas relacionadas con el procesamiento de imágenes, especialmente porque están diseñadas para reconocer patrones directamente desde imágenes pixeladas con un mínimo preprocesamiento (LeCun *et al.*, 2015). Esta capacidad de las CNN se inspira en la biología, tomando como referencia la forma en que se cree que el córtex visual del cerebro humano procesa la información visual.

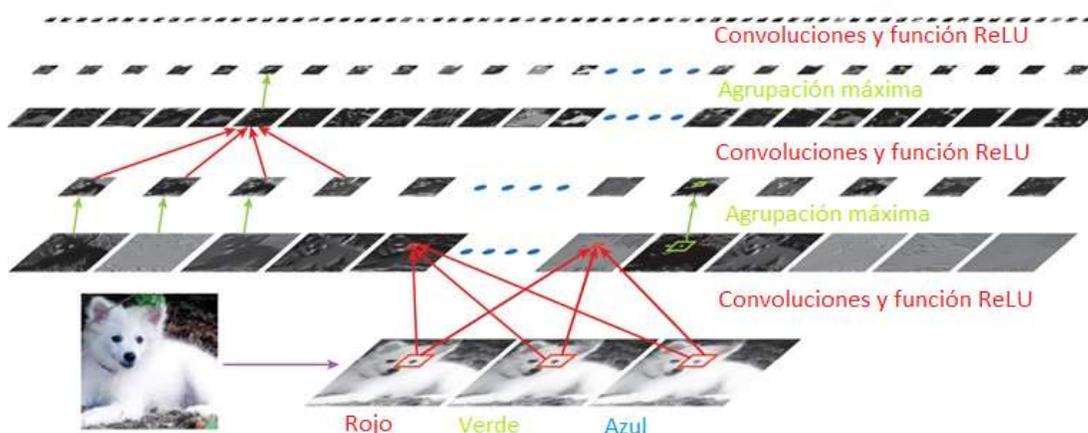
La capa convolucional se destaca como el componente central del diseño de una CNN. Esta capa se encarga de aplicar una serie de filtros sobre la imagen para extraer detalles cruciales como bordes, texturas y colores mediante un proceso conocido como convolución. Después de este proceso, se aplica una función de activación, comúnmente la función ReLU (Rectified Linear Unit), que introduce no linealidad al modelo, permitiendo que la red capture relaciones más complejas en los datos.

Posteriormente, el proceso de pooling o submuestreo entra en juego. Su principal función es reducir las dimensiones espaciales de la imagen, es decir, su ancho y alto, pero conservando las características más significativas. El método "Max Pooling" elige el valor máximo de una región específica de la imagen. Esta es una de las estrategias más utilizadas durante esta fase.

Una vez que la red ha pasado por varias capas convolucionales y de pooling, se llega a las capas totalmente conectadas. Estas capas, similares a las de una red neuronal tradicional, se encargan de realizar la clasificación final basada en las características previamente extraídas. Finalmente, en la etapa de salida, la CNN utiliza una función Softmax para convertir las salidas en probabilidades, y durante el entrenamiento, se emplea una función de pérdida para ajustar los pesos de la red en función de la diferencia entre la salida predicha y la real tal como se muestra en la Figura 11.

Figura 11

Proceso de una red neuronal convolucional (CNN)



Nota. Representa el proceso de aprendizaje que realiza una red neuronal convolucional entrenada para reconocer perros, donde dada una imagen se comienza a extraer las características distintivas del animal y se aplican iteraciones (convoluciones) en las cuales se maximiza la precisión del reconocimiento y se trata de minimizar el error. Tomado de *Deep learning* (p. 438), por LeCun *et al.*, 2015, Nature, 521 (7553).

Las CNN han demostrado ser herramientas cruciales en el campo de la Visión por Computadora. Esto se debe a su capacidad para aprender características jerárquicas de las imágenes, resultando en ser adecuadas para tareas como la clasificación de imágenes, el reconocimiento facial, la detección de objetos y la segmentación semántica (LeCun *et al.*, 2015).

Sin embargo, es fundamental reconocer que, aunque las CNN ofrecen múltiples ventajas, como el Machine Learning de características y su eficiencia en términos de parámetros, también presentan desafíos. Por ejemplo, el entrenamiento de una CNN puede requerir grandes conjuntos de datos etiquetados y ser computacionalmente intensivo.

Gait Energy Image (GEI)

El análisis de la forma de caminar, conocido como "gait", ha emergido como una característica biométrica relevante en el campo de la visión por computadora y la identificación de personas (Bakchy *et al.*, 2022). Una de las técnicas más destacadas en este ámbito es la Imagen de Energía de la Marcha, o GEI (por sus siglas en inglés, Gait Energy Image). El GEI es una representación espacial que captura la energía de movimiento de una persona mientras camina, y se ha demostrado que es efectiva para la identificación de individuos basada en su forma de caminar (Han & Bhanu, 2006; Mogan *et al.*, 2022).

El GEI se construye promediando las siluetas binarias de una persona a lo largo de una secuencia completa de caminata. Esta secuencia captura el movimiento y la forma del cuerpo durante diferentes fases de la marcha, y al promediar estas siluetas, se obtiene una imagen que encapsula la energía de la marcha de la persona (ver Figura 12). Esta imagen resultante, el GEI, sirve como una representación distintiva de la forma de caminar del individuo.

Figura 12

Representación de siluetas en una secuencia de caminata



Nota. Representa la silueta de una persona en diferentes frames, donde se aprecia las variaciones conforme ocurre la marcha o caminata. Tomado de *Individual recognition using gait energy image* (p. 317), por Han & Bhanu, 2006, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28 (2).

Una de las principales ventajas del GEI es su resistencia a las variaciones temporales y las pequeñas diferencias en las posturas de caminata entre diferentes secuencias. Además, al ser una

representación espacial, el GEI puede ser procesado utilizando técnicas convencionales de procesamiento de imágenes, lo que facilita su integración en sistemas de visión por computadora.

En el contexto de la Re-Identificación de personas, el GEI ofrece un enfoque no intrusivo para identificar individuos basándose en una característica que es difícil de ocultar o falsificar: la forma de caminar. Aunque la técnica presenta ventajas significativas, también enfrenta desafíos, especialmente en escenarios donde las condiciones de iluminación varían o cuando hay obstrucciones en la vista de la persona que camina.

Histogram of Oriented Gradients (HoG)

En el ámbito de la visión por computadora, la capacidad de extraer características distintivas de las imágenes es esencial para diversas tareas, desde la detección de objetos hasta la Re-Identificación de personas. Los Histogram of Oriented Gradients, o HoG, es un método que se ha popularizado en este ámbito, el cual es un descriptor de características que se ha demostrado particularmente efectivo para la detección de objetos, especialmente en aplicaciones como la detección de peatones (Dalal & Triggs, 2005b).

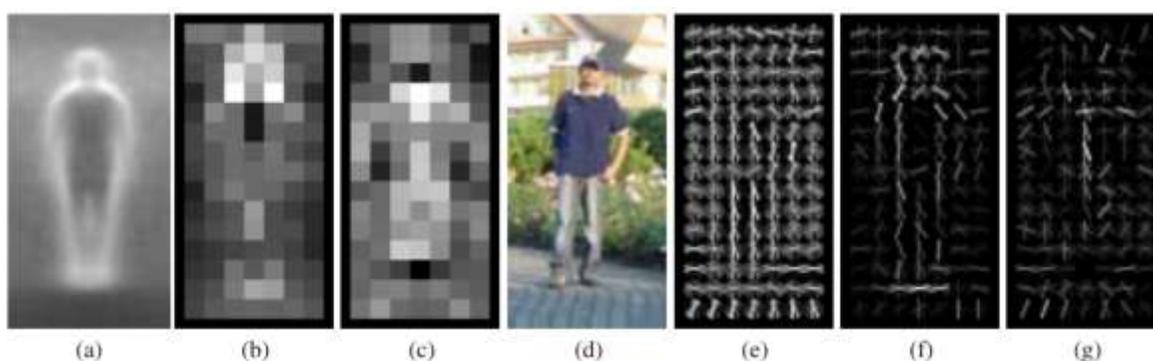
El principio subyacente del HoG es bastante intuitivo. La técnica se basa en la premisa de que las formas y apariencias locales de un objeto en una imagen pueden ser descritas por la distribución de las direcciones de sus bordes o gradientes. Por lo tanto, HoG captura la intensidad y orientación de los gradientes en regiones locales de una imagen (Tyagi, 2021).

El proceso para calcular el HOG comienza con la aplicación de un filtro de gradiente a la imagen para obtener componentes de gradiente en las direcciones x e y. Estos componentes se utilizan para calcular la magnitud y dirección del gradiente para cada píxel. Luego, la imagen se divide en pequeñas celdas, y para cada celda, se crea un histograma de gradientes orientados. Estos histogramas se normalizan, y la concatenación de estos histogramas normalizados forma el descriptor HOG de la imagen.

Una de las fortalezas del HoG radica en su capacidad para capturar información de forma y estructura, lo que lo hace resistente a variaciones en iluminación y sombras. Esta resistencia es esencial en aplicaciones de detección en condiciones de iluminación variables. Sin embargo, como cualquier técnica, HoG tiene sus limitaciones, especialmente cuando se enfrenta a cambios significativos en la postura o apariencia del objeto de interés.

Figura 13

Descriptor de características HOG



Nota. Muestra que los detectores HoG se centran principalmente en los contornos de la silueta corporal, donde a) Muestra la imagen promedio del gradiente en los ejemplos de entrenamiento, b) Cada "píxel" muestra el peso máximo positivo de SVM en el bloque centrado en el píxel, c) De manera similar, muestra los pesos negativos de SVM, d) Presenta una imagen de prueba, e) Muestra el descriptor HOG calculado de la imagen, f) El descriptor HoG ponderado por los pesos positivos de SVM, g) El descriptor R-HoG ponderado por los pesos negativos de SVM. Tomado de *Histograms of oriented gradients for human detection* (p. 886), por Dalal & Triggs, 2005, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 1.

En el presente trabajo HoG es una herramienta valiosa para extraer características distintivas de las imágenes, especialmente cuando se trata de características softbiométricas como la silueta corporal. Su capacidad para capturar detalles estructurales lo hace adecuado para identificar y diferenciar entre diferentes individuos basándose en sus características visuales.

You Only Look Once (YOLO)

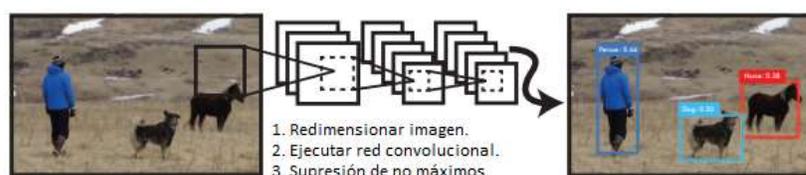
YOLO, cuyo nombre proviene del acrónimo "You Only Look Once" (Solo Miras Una Vez), es una técnica innovadora en el campo de la Visión por Computadora, particularmente útil en tareas de detección de objetos. A diferencia de los enfoques tradicionales que segmentan una imagen en múltiples regiones y luego ejecutan la detección de objetos en cada una de ellas, YOLO adopta un enfoque completamente diferente. Analiza la imagen en su totalidad y predice los objetos y sus ubicaciones en un solo paso, lo que le permite ser extremadamente rápido y eficiente (Redmon *et al.*, 2016b).

El funcionamiento de YOLO se basa en dividir la imagen en una cuadrícula de tamaño determinado, por ejemplo, 13x13. Para cada celda de esta cuadrícula, el modelo predice múltiples bounding boxes (cajas delimitadoras) y sus respectivas confianzas. Estas confianzas reflejan la probabilidad de que la caja contenga un objeto y qué tan bien se ajusta la caja al objeto.

Además, para cada celda, el modelo también predice las probabilidades de clase para cada objeto detectado. Estas probabilidades indican la probabilidad de que el objeto pertenezca a una clase específica, como "persona", "automóvil" o "perro", una ejemplificación del funcionamiento de YOLO se muestra en la Figura 14.

Figura 14

Representación de YOLO



Nota. Muestra el sistema de detección de objetos YOLO, donde en 1) cambia la imagen de tamaño, 2) esta imagen pasa por la arquitectura de red YOLO, 3) Se empareja el resultado de la imagen con los conjuntos de datos de entrenamiento para mostrar la clasificación de este objeto sea: persona o animal.

Tomado de *You Only Look Once: Unified, Real-Time Object Detection* (p. 779), por Radmon *et al.*, 2016, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Así también, YOLO resulta computacionalmente viable debido a su velocidad. Dado que solo "mira" la imagen una vez, puede detectar objetos en tiempo real, lo que lo hace ideal para aplicaciones en tiempo real, como la detección de objetos en vehículos autónomos o sistemas de vigilancia en vivo. Sin embargo, esta velocidad viene con un costo: YOLO tiende a tener más falsos positivos en comparación con otros detectores más lentos, pero más precisos.

Otro aspecto distintivo de YOLO es su capacidad para detectar objetos en contextos variados. A diferencia de otros modelos que se centran en partes individuales de una imagen, YOLO tiene en cuenta el contexto global, lo que le permite reconocer objetos incluso en situaciones ambiguas o en presencia de oclusiones.

Proceso de Re-Identificación y Fusión de Modelos

La Re-Identificación de personas por medio de la combinación de características biométricas y soft-biométricas ha demostrado ser una estrategia efectiva para mejorar la precisión en la identificación, alcanzando tasas de reconocimiento de hasta el 95% en ciertos escenarios controlados (Zheng *et al.*, 2016). Esta mejora sustancial se debe a la incorporación de las características Soft-biométricas, las cuales brindan patrones que, aunque no sean únicos para cada persona, permiten incrementar la precisión de los modelos de reconocimiento.

Re-Identificación facial y forma de caminar

La identificación basada en el rostro y la forma de caminar ha mostrado ser altamente efectiva. Utilizando técnicas como Haar Cascade y CNN para la detección facial, y GEI para la identificación basada en la forma de caminar, investigaciones recientes han reportado tasas de acierto de hasta el 92% en ambientes controlados (Mostafaie *et al.*, 2020).

La Re-Identificación facial y de la forma de andar es un procedimiento que integra una serie de modelos y enfoques de Visión por Computador para identificar de forma fiable a las personas tanto en contextos controlados como no controlados.

Inicialmente, las cámaras de seguridad capturan imágenes o secuencias de video de los individuos. Estos datos son preprocesados para mejorar la calidad, eliminando ruido, ajustando el contraste y normalizando la iluminación. Este paso es esencial para garantizar que las técnicas posteriores operen en condiciones óptimas (Zhang *et al.*, 2010).

Para la identificación facial, se utiliza el algoritmo Haar Cascade, también conocido como Viola-Jones. Al detectar patrones concretos de rasgos faciales, como los ojos, la nariz y la boca, este sistema localiza rápidamente rostros en las fotos. Una vez detectada la cara, se extrae y se aísla para su posterior análisis (Viola & Jones, 2001).

Para extraer los rasgos distintivos de un rostro, se utilizan redes neuronales convolucionales (CNN). Estas redes son capaces de identificar y aprender patrones complejos en las imágenes, lo que las hace ideales para el reconocimiento facial. Las características obtenidas son codificadas en un vector que representa de manera única el rostro de la persona (Lawrence *et al.*, 1997).

Para el reconocimiento por la forma de caminar se debe identificar y extraer la silueta del individuo mientras este camina, para ello se utiliza YOLO. Esta técnica de detección de objetos en tiempo real es eficiente para identificar la figura humana en secuencias de video.

Ahora bien, obtenidas las siluetas se procede a usar Gait Energy Image (GEI) la cual es una técnica que captura la forma única en que una persona camina (Bakchy *et al.*, 2022). Se crea una representación visual de la forma de caminar al promediar las siluetas binarias de una secuencia de caminata. Esta representación captura la dinámica y el movimiento del individuo, permitiendo su identificación basada en su forma de caminar (Han & Bhanu, 2006).

Una vez extraídas y codificadas las características faciales y de la forma de caminar, se comparan con una base de datos de características previamente almacenadas. Utilizando técnicas de comparación y métricas de similitud, se determina la correspondencia más cercana, logrando así la Re-Identificación del individuo.

Finalmente, los resultados se validan con datos etiquetados para asegurar la precisión del sistema. Si se detectan errores o imprecisiones, el sistema se retroalimenta para mejorar en futuras identificaciones.

Re-Identificación basada en color, textura y silueta corporal

Las características Soft-biométricas (silueta corporal, color y textura de la vestimenta) han demostrado ser complementarias a las características Biométricas. Al emplear técnicas como YOLO y HOG para extraer estas características, se ha logrado mejorar la precisión de la Re-Identificación en un 8% adicional en comparación con el uso exclusivo de características Biométricas (R. Zhao *et al.*, 2013).

Aunque estas características no son únicas para cada persona, cuando se combinan y se procesan con técnicas avanzadas de Visión por Computadora, pueden ofrecer un reconocimiento efectivo en ambientes controlados.

Inicialmente, las cámaras de seguridad capturan imágenes o secuencias de video de los sujetos en movimiento. Estos datos brutos son sometidos a un proceso de preprocesamiento para mejorar su calidad, eliminando ruido, ajustando el contraste y equilibrando la iluminación, lo que facilita la posterior extracción de características (Ojala *et al.*, 2002b).

Para identificar y segmentar la figura humana de otros elementos en la escena, se utiliza YOLO (You Only Look Once). Esta técnica permite detectar y obtener la forma del cuerpo humano mientras se procesa en tiempo real una entrada de video procedente de una cámara u otro medio que capture secuencias de imágenes (Redmon *et al.*, 2016b).

El análisis del color se centra en identificar y extraer información sobre los colores predominantes en la vestimenta y piel del individuo. Esta información es crucial, ya que el color de la ropa o el tono de piel pueden ser indicativos y distintivos en ciertos contextos (Zheng *et al.*, 2016).

El Patrón Binario Local (LBP) se aplica para discernir y extraer características de textura de la vestimenta del individuo. La textura, que puede manifestarse como patrones específicos en la ropa o en la piel, como lunares o cicatrices, es analizada para proporcionar datos adicionales para la Re-Identificación (Ojala *et al.*, 2002b).

Las redes neuronales convolucionales (CNN) se emplean para un análisis más detallado de la silueta corporal. Estas redes son capaces de identificar y aprender patrones complejos en las imágenes, lo que las hace ideales para este propósito (LeCun *et al.*, 2015).

Con las características de color, textura y silueta extraídas y codificadas, se procede a compararlas con una base de datos previamente establecida. Utilizando algoritmos avanzados y métricas de similitud, se busca la correspondencia más cercana, logrando así la Re-Identificación del individuo (Zheng *et al.*, 2016).

Los resultados obtenidos son validados contra un conjunto de datos etiquetados para asegurar la precisión del sistema. Si se identifican áreas de mejora, el sistema se ajusta y se retroalimenta para optimizar el rendimiento en futuras identificaciones.

Proceso de fusión de características y modelos

La fusión de características y modelos es un paso crucial para maximizar la precisión en la Re-Identificación. Al combinar múltiples características y modelos, investigaciones recientes han logrado tasas de reconocimiento de hasta el 97%, evidenciando la eficacia de este enfoque integrador (Chen *et al.*, 2017).

Inicialmente, se extraen características de cada modalidad por separado. Por ejemplo, del rostro se extraen características utilizando Haar Cascade y CNN, mientras que de la forma de caminar se extrae

características mediante GEI y YOLO. Simultáneamente, de la vestimenta y la silueta corporal se extraen características de color, textura y forma utilizando técnicas como LBP y HOG (Dalal & Triggs, 2005b).

Con las características extraídas, se generan modelos independientes para cada modalidad. Estos modelos son entrenados utilizando conjuntos de datos específicos para cada modalidad y validados para asegurar su precisión (LeCun *et al.*, 2015).

Seguidamente, las características extraídas de diferentes modalidades se combinan para formar un vector de características unificado. Este vector combina información biométrica y softbiométrica, proporcionando una representación más completa del individuo (Zheng *et al.*, 2016).

El proceso de entrenamiento para el modelo fusionado utiliza un conjunto de datos que incorpora tanto datos biométricos como datos soft-biométricos a nivel de características (Hua *et al.*, 2022). Este entrenamiento busca optimizar el modelo para que pueda identificar individuos con precisión incluso cuando alguna de las modalidades no esté claramente visible o sea ambigua (Zheng *et al.*, 2016).

Finalmente, el modelo fusionado se valida contra un conjunto de datos etiquetados para asegurar su precisión. Si se identifican áreas de mejora, el modelo se ajusta y se retroalimenta para optimizar su rendimiento en futuras identificaciones.

Tecnologías de implementación

Para la implementación del sistema de Re-Identificación de personas mediante características Biométricas y Soft-biométricas se hizo uso de herramientas y tecnologías ampliamente utilizadas en la industria, esto garantiza que se puedan obtener resultados similares a otras implementaciones realizadas en investigaciones previas relacionadas a nuestro tema de investigación.

Python es un lenguaje de programación de alto nivel que ha ganado popularidad en los campos de la Visión por Computadora y el Machine Learning por su sintaxis sencilla y su amplia biblioteca de herramientas especializadas (Van Rossum & Drake, 2009). Su naturaleza interpretada y su capacidad

para integrarse con otras tecnologías lo hacen ideal para el desarrollo y la implementación de algoritmos complejos en este campo.

La implementación de métodos como Haar Cascade y la extracción de rasgos faciales emplean OpenCV, una biblioteca de código abierto con experiencia en procesamiento de imágenes y Visión por Computador. OpenCV, que se creó inicialmente en C/C++, cuenta con una interfaz completa en Python, lo que facilita su incorporación y uso en proyectos basados en Python.

TensorFlow, desarrollado por Google, es una plataforma de código abierto para el Machine Learning y las redes neuronales profundas. Su flexibilidad y capacidad para implementar y entrenar modelos como GEI, YOLO y CNN lo hacen indispensable. Además, su API en Python permite una integración fluida y un desarrollo más rápido (Abadi *et al.*, 2016).

Scikit-learn es otra biblioteca esencial escrita en Python que proporciona herramientas para el análisis de datos y el Machine Learning. Es especialmente útil para la implementación de algoritmos de clasificación como SVM (Pedregosa *et al.*, 2011).

Métricas de evaluación

Es fundamental comprender qué métricas de evaluación se utilizan para medir la eficacia y la precisión de los modelos creados. En concreto, para evaluar el rendimiento se utilizan medidas de exactitud, precisión y recall.

La métrica de Accuracy (Exactitud) se utiliza para medir la proporción de muestras correctamente clasificadas sobre el total de muestras evaluadas. Este valor se da en forma de porcentaje y sirve como indicador general del rendimiento del modelo. La métrica de precisión se utiliza para evaluar la capacidad del modelo para re-identificar con éxito a los individuos utilizando la información biométrica y soft-biométrica obtenida.

De todas las muestras clasificadas como positivas, las métricas de precisión informan del número de muestras positivas detectadas correctamente (Vishwakarma & Upadhyay, 2018). Las

medidas de precisión se utilizan para comprobar la precisión de los modelos que clasifican a las personas como positivas. Una alta precisión significa que el modelo detecta correctamente a las personas que son verdaderos positivos y tiene una baja tasa de falsos positivos.

El indicador Recall se utiliza para calcular el porcentaje de muestras que son realmente positivas (Nixon *et al.*, 2006). Recall se utiliza para medir la capacidad de un modelo para identificar a cada individuo que es un verdadero positivo. Un valor Recall alto significa que el modelo identifica correctamente a todos los individuos que hay que identificar y tiene un bajo índice de falsos negativos.

La exactitud, la precisión y el recall proporcionan una evaluación detallada del rendimiento global del modelo de Re-Identificación de personas. Estas métricas proporcionan una imagen más completa de la capacidad del sistema para identificar y clasificar personas con precisión. Para un modelo generado, una estrategia equilibrada que produzca valores altos para las tres métricas se considera un desempeño satisfactorio.

Tabla 1

Métricas de evaluación utilizadas

Métrica de evaluación	Formula aplicada
Accuracy	$accuracy = \frac{VP + VN}{VP + VN + FP + FN}$
Precision	$precision = \frac{VP}{VP + FP}$
Recall	$recall = \frac{VP}{VP + FN}$

En la Tabla 1: Métricas de evaluación utilizadas, se presentan las fórmulas matemáticas que permiten la evaluación cuantitativa de los resultados del sistema de Re-Identificación de personas propuesto, a continuación, se menciona la interpretación de las variables que intervienen en las fórmulas.

- **VP:** Representa el número de imágenes positivas que resultaron en una clasificación correcta de verdaderos.
- **VN:** Representa el número de imágenes negativas que resultaron en una clasificación correcta de falsos.
- **FN:** Representa el número de imágenes positivas que resultaron en una clasificación incorrecta de falsos.
- **FP:** Representa el número de imágenes negativas que resultaron en una clasificación incorrecta de positivos.

Tabla 2

Valores de la matriz de confusión

	Positivo	Negativo
Positivo	Verdadero Positivo (VP)	Falso Negativo (FN)
Negativo	Falso Positivo (FP)	Verdadero Negativo (VN)

Así también, en la Tabla 2: Valores de la matriz de confusión, detalla los casos posibles cuando el modelo evalúa una imagen, estos son:

- **1:** Si re-identifica a la persona.
- **0:** No re-identifica a la persona.

Entonces, con los valores descritos los resultados obtenidos pueden ser interpretados como:

Verdadero Positivo: Corresponde al valor (1,1), es decir 1.

Falso Negativo: Corresponde al valor (1,0), es decir 0.

Falso Positivo: Corresponde al valor (0,1), es decir 1.

Verdadero Negativo: Corresponde al valor (0,0), es decir 0.

Marco de trabajo

La técnica Scrum, conocida por su enfoque ágil y adaptable, se ha convertido en uno de los principales métodos de gestión de proyectos de desarrollo de software (Rola & Kuchta, 2015). Esta metodología se basa en principios que promueven la colaboración entre equipos multidisciplinarios, la adaptabilidad ante cambios y la entrega continua de valor.

En el presente trabajo se utilizó Scrum por la complejidad (desafíos técnicos) y la incertidumbre del proyecto. Dado que la Visión por computadora y las técnicas de Re-Identificación están en constante evolución, fue importante escoger un enfoque ágil y flexible que permita adaptarse rápidamente a los cambios.

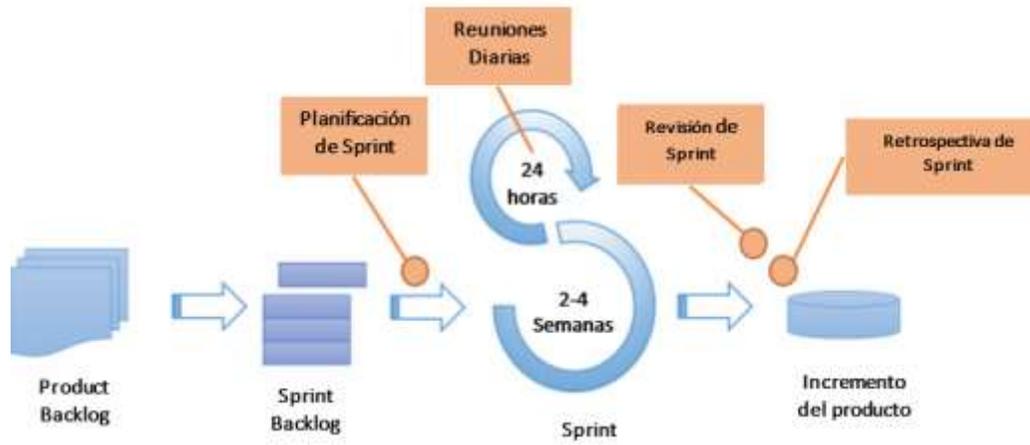
Scrum se organiza en ciclos llamados "sprints", que suelen durar entre dos y cuatro semanas. El equipo se reúne para planificar las tareas que deben llevarse a cabo al inicio de cada sprint utilizando un "backlog" o lista priorizada de necesidades (Navajo, 2023). El equipo colabora para completar estas tareas a lo largo del sprint y, al final del mismo, se proporciona un incremento del producto, es decir, una versión funcional del sistema o software (Navajo, 2023; Rola & Kuchta, 2015).

Una de las principales ventajas de Scrum es la retroalimentación. Al final de cada sprint, el equipo realiza una revisión y una retrospectiva, donde se evalúa el trabajo realizado y se identifican áreas de mejora (Navajo, 2023). Esto permite adaptarse y mejorar continuamente, garantizando que el proyecto se mantenga alineado con los objetivos planteados.

Para el presente trabajo, Scrum ofrece un marco estructurado que facilita la colaboración y comunicación (Navajo, 2023). Además, al adoptar un enfoque iterativo, se pueden realizar pruebas y ajustes continuos en las técnicas y modelos implementados, asegurando que el sistema de Re-Identificación sea lo más preciso y eficiente posible. En la Figura 15 se presenta el proceso del marco de trabajo scrum.

Figura 15

Ciclo de vida de Scrum



Nota. Tomado de (Zayat & Senvar, 2020)

Capítulo III

Proceso de implementación del sistema de Re-Identificación de personas

En el ámbito de la visión por ordenador, la construcción de un sistema de Re-Identificación humana basado en rasgos biométricos y soft-biométricos plantea un desafío significativo. Este proceso no solo implica la aplicación de técnicas y modelos avanzados, sino también la integración efectiva de múltiples componentes para lograr un sistema cohesivo y eficiente. La relevancia de una implementación adecuada radica en su capacidad para traducir teorías y conceptos en una solución práctica y aplicable en escenarios del mundo real (Merusi, 1995).

Dentro del contexto de esta investigación, el proceso de implementación se llevó a cabo siguiendo la metodología Scrum (ver Figura 15), una de las metodologías ágiles más reconocidas y utilizadas en el desarrollo de software. Scrum, con su enfoque iterativo e incremental, permite adaptarse a los cambios y requerimientos emergentes, garantizando que el producto final cumpla con las expectativas y necesidades del proyecto (Rola & Kuchta, 2015).

Este capítulo tiene como objetivo detallar el proceso de implementación del sistema, desde la definición de historias de usuario y la creación del Product Backlog, hasta las validaciones y métricas de evaluación utilizadas. A través de este recorrido, se busca ofrecer una visión completa y estructurada de cómo se llevó a cabo el desarrollo del sistema, resaltando las decisiones tomadas, los desafíos enfrentados y las soluciones implementadas.

Descripción del equipo Scrum

Para el desarrollo se establecieron roles que deberán desempeñar cada uno de los miembros del equipo, y determinar sus responsabilidades y funciones. En la Tabla 3, se muestran el rol, integrante y sus funciones.

Tabla 3

Equipo Scrum

N°.	Rol Scrum	Integrante	Funciones
01	Product Owner	Dr. José Luis Carrillo Medina	Responsable de especificar los requerimientos del sistema, definir los tiempos de entrega y evaluar los resultados.
02	Scrum Master	Bryan Stiven Chicaiza Chicaiza	Encargado del cumplimiento de las fechas establecidas en cada Sprint, es el líder del equipo.
03	Development Team	Lenin Sebastián Arteaga Chapi Bryan Stiven Chicaiza Chicaiza	Encargados del desarrollo del sistema de Re-Identificación de personas.

Nota. Estructura del Equipo Scrum para el desarrollo del sistema de Re-Identificación de personas.

De acuerdo con los roles presentados en la Tabla 3, se cubren roles necesarios para la implementación de Scrum como metodología de trabajo, así también por la limitación de integrantes algunos de ellos deben cumplir 2 roles a la vez, a fin de cumplir las actividades de las historias de usuario en los plazos establecidos.

Historias de Usuario

En la Tabla 4 se presentan las historias de usuario correspondientes al sistema de Re-Identificación de personas, cada historia cuenta con un identificador único (ID), nombre, rol, funcionalidad y resultado.

Tabla 4

Historias de Usuario

ID	Nombre	Rol	Funcionalidad	Resultado
01	H.U.01	Usuario	Quiero un dataset de personas del cual se pueda extraer características biométricas y softbiométricas.	Para la extracción de características biométricas (rostro, forma de caminar) y softbiométricas (silueta corporal, color y textura de la vestimenta).
02	H.U.02	Usuario	Quiero un modelo y/o algoritmo de Visión por Computador que permita reconocer una persona mediante su rostro mientras se procesa la entrada de video de la primera cámara de seguridad.	Para identificar una persona mediante el rostro.

ID	Nombre	Rol	Funcionalidad	Resultado
03	H.U.03	Usuario	Quiero un modelo y/o algoritmo de Visión por Computador que permita reconocer una persona mediante su forma de caminar mientras se procesa la entrada de video de la primera cámara de seguridad.	Para identificar una persona mediante su forma de caminar.
04	H.U.04	Usuario	Quiero un modelo y/o algoritmo de Visión por Computadora que permita identificar una persona mediante características softbiométricas	Para identificar una persona mediante la fusión de características y softbiométricas (silueta corporal, color y textura de la vestimenta).

ID	Nombre	Rol	Funcionalidad	Resultado
			(silueta corporal, color y textura de la vestimenta) mientras se procesa la entrada de video de la primera cámara de seguridad.	
05	H.U.05	Usuario	Quiero un modelo y/o algoritmo de Visión por Computadora que permita identificar una persona mediante características biométricas (rostro y forma de caminar) y soft-biométricas (silueta corporal, color y textura de la vestimenta).	Para validar que el sistema identifica a la persona mediante características biométricas (rostro, forma de caminar) y soft-biométricas (silueta corporal, color y textura de la vestimenta).

ID	Nombre	Rol	Funcionalidad	Resultado
			la vestimenta) mediante la fusión de modelos y/o algoritmos de Re-Identificación.	
06	H.U.06	Usuario	Quiero que el sistema de Re- Identificación permita el reconocimiento de una persona a través de características biométricas (rostro y forma de caminar) y softbiométricas (silueta corporal, color y textura de la vestimenta) mientras se procesa la primera entrada	Para validar que el sistema generé los modelos de reconocimiento con la primera cámara de seguridad y re- identifiqué a la persona con la segunda cámara de seguridad.

ID	Nombre	Rol	Funcionalidad	Resultado
			de video de la primera cámara de seguridad y se Re-Identifiqué en tiempo real a la persona mientras se procesa la segunda entrada de video procedente de la segunda cámara de seguridad.	

Product Backlog del proyecto

Después de definir las historias de usuario, se procede a definir el Product Backlog es una lista de tareas priorizadas que se llevaran a cabo durante un sprint (Navajo, 2023; Rola & Kuchta, 2015). El Product Backlog está compuesto por un conjunto de historias de usuario donde se detallan los requisitos del producto desde el punto de vista del usuario (Navajo, 2023).

Cada elemento de Product Backlog se ordena en base a su prioridad para el usuario final, es decir el valor que cada elemento aporta, esto permite al equipo Scrum desarrollar de manera ágil aquellas tareas que impliquen las funcionalidades más importantes.

En la Tabla 5, se listan las historias de usuario correspondientes al sistema de Re-Identificación de personas con: estimación de tiempo, fechas de inicio y finalización, así como el sprint al cual pertenecen, estos elementos están priorizados por nivel de importancia.

Tabla 5*Product Backlog del Proyecto*

Historia de Usuario	Nombre	Estimación (días)	Fecha inicio	Fecha fin	N° de Sprint
1	H.U.01	7	2/05/2023	10/05/2023	01
2	H.U.02	10	11/05/2023	24/05/2023	02
3	H.U.03	12	25/05/2023	9/06/2023	03
4	H.U.04	13	12/06/2023	28/06/2023	04
5	H.U.05	10	29/06/2023	12/07/2023	05
6	H.U.06	25	13/07/2023	18/08/2023	06

En la Tabla 6: Distribución de los Sprint del proyecto, se listan cada uno de los sprint que comprenden al sistema de Re-Identificación de personas, cada elemento cuenta con: estimación de tiempo, fechas de inicio y finalización, donde cabe mencionar que el comienzo del proyecto fue el 2 de mayo de 2023 y su finalización el 18 de agosto de 2023.

Así también, la distribución elegida fue de lunes a viernes con una jornada laboral de 8 horas diarias, además la estimación de la duración de cada Sprint, fue considerada en base a la experiencia y dificultad de cada tema tratado por sprint.

Tabla 6*Distribución de los Sprint del proyecto*

N° de Sprint	Nombre	Estimación (días)	Estimación (horas)	Jornada (horas)	Fecha inicio	Fecha fin
01	Inicio y Preparación	7	56	8	2/05/2023	10/05/2023

N° de Sprint	Nombre	Estimación (días)	Estimación (horas)	Jornada (horas)	Fecha inicio	Fecha fin
02	Desarrollo de Modelo Facial	10	80	8	11/05/2023	24/05/2023
03	Desarrollo de Modelo Forma de Caminar	12	96	8	25/05/2023	9/06/2023
04	Desarrollo de Modelos Soft-biométricos	13	104	8	12/06/2023	28/06/2023
05	Fusión de Modelos	10	80	8	29/06/2023	12/07/2023
06	Validación, Ajustes de Modelos e integración	25	200	8	13/07/2023	18/08/2023

Ambiente para la implementación del sistema de Re-Identificación de personas

La implementación del sistema de Re-Identificación de personas se efectúa en un ambiente controlado, el cual está ubicado en la Universidad de las Fuerzas Armadas ESPE Sede Latacunga, de los cuales se ha escogido uno de los laboratorios de la institución del Departamento de Ciencias de la Computación. Las cámaras de seguridad están colocadas de manera contigua, donde la primera cámara tiene su punto de vista (FOV) norte, mientras que la segunda tiene su punto de vista (FOV) el costado.

Cabe mencionar que en este espacio no existen variaciones de iluminación, ni oclusión de objetos o personas, en este entorno es donde se realizan las pruebas para el sistema de Re-Identificación, como se ve en la Figura 16 y Figura 17.

Figura 16

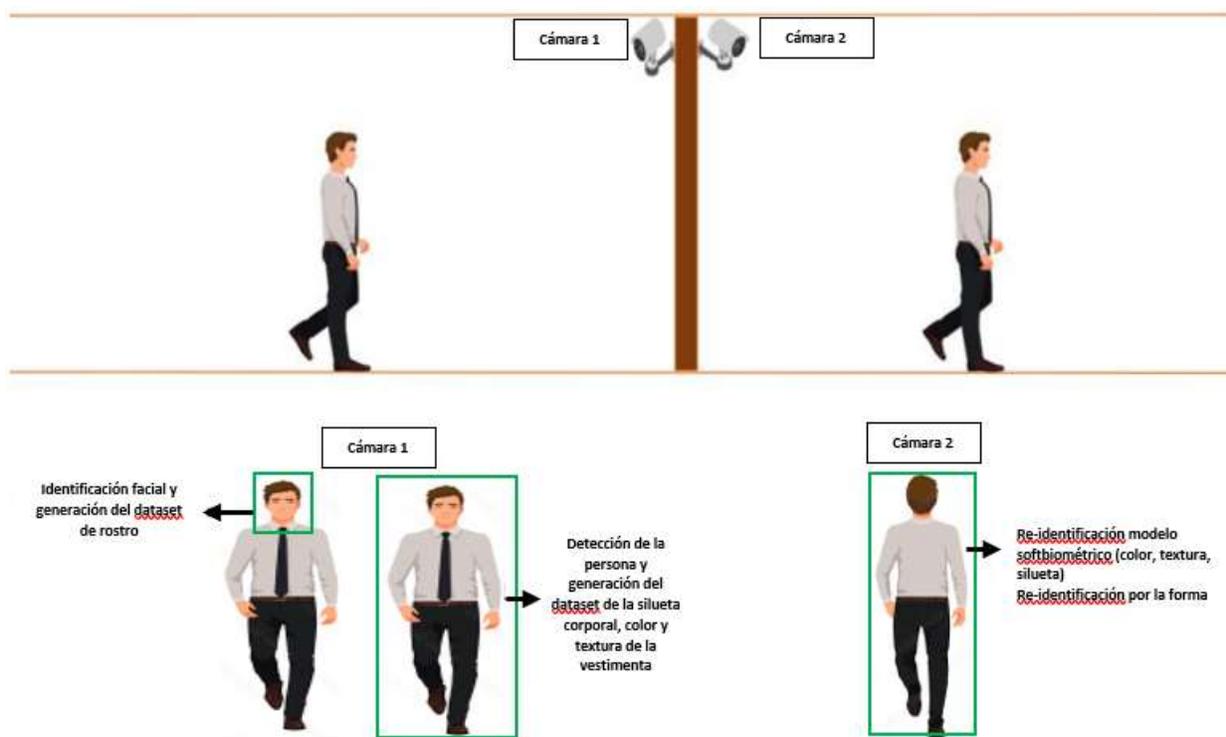
Entorno controlado para pruebas del sistema de Re-Identificación de personas



Nota. Las pruebas al sistema de Re-Identificación de personas se realizan en los pasillos de la Universidad de las Fuerzas Armadas Espe Sede Latacunga.

Figura 17

Ubicación de las cámaras utilizadas para capturar los videos de prueba



Nota. La información capturada con la primera cámara se identifica el rostro y detecta a la persona.

Estos datos son utilizados para generar los datasets empleados para entrenar los modelos y algoritmos de Machine Learning utilizados (Haar Cascade, SVM, CNN), mientras que la segunda cámara se utiliza para Re-identificar a la persona.

Sprint 1: Inicio y Preparación

La base de cualquier sistema de visión por computadora es el conjunto de datos con el que se entrena. En este caso, se requiere un dataset que contenga imágenes y videos de individuos desde diferentes ángulos, con variaciones en la iluminación y en diferentes posturas. Este dataset será la piedra angular para desarrollar modelos precisos y eficientes.

Historia de Usuario 01: Creación de Dataset que contenga características biométricas y softbiométricas.

La historia de usuario H.U.01 indica que se requiere un Dataset de personas del cual se pueda extraer características biométricas y softbiométricas, con estas características se procede al entrenamiento de los modelos que utilizan características biométricas (rostro y forma de caminar) y características softbiométricas (silueta corporal, color y textura de la vestimenta).

Detalle Historia de Usuario 01. Esta historia de usuario consiste en establecer las bases para el desarrollo del proyecto, asegurando que se cuente con todos los recursos y datos necesarios para iniciar el proceso de implementación.

Tabla 7

Historia de usuario para la obtención del dataset

Historias de Usuario	
Numero: H.U.01	Usuario: Usuario
Nombre historia: Obtención de Datasets	Numero Sprint: 1
Prioridad: Alta	Riesgo de desarrollo: Baja
Puntos estimados (días): 7	interacción asignada: 1
Desarrolladores responsables: Bryan Chicaiza, Lenin Arteaga	
Descripción: Como usuario quiero un dataset de personas del cual se pueda extraer características biométricas y softbiométricas.	
Validación:	
<ul style="list-style-type: none"> • Se obtiene imágenes de personas en el entorno controlado. • Se asignan rutas de almacenamiento únicas para las imágenes de cada persona. 	

Sprint Backlog 01

En el contexto de nuestra investigación sobre la Re-Identificación de personas, el Sprint 1, que se llevó a cabo del 2 al 10 de mayo de 2023, se centró en la fase inicial y de preparación del proyecto.

La Tabla 8 detallan las tareas que se realizaron durante el Sprint 1.

Tabla 8

Sprint Backlog 01

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.01	Investigación en bases de datos académicas y repositorios especializados.	16	2/05/2023	3/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.01	Selección de datasets que contengan imágenes y videos de rostros, formas de caminar, y características softbiométricas.	16	4/05/2023	5/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.01	Descarga y organización de los datos en	16	8/05/2023	9/05/2023	Bryan Chicaiza,	Finalizado

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	carpetas y subcarpetas según tipo de dato (rostro, forma de caminar, silueta corporal, color y textura de la vestimenta).				Lenin	
					Arteaga	
H.U.01	División de los datasets en conjuntos de entrenamiento, validación y prueba.	8	10/05/2023	10/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

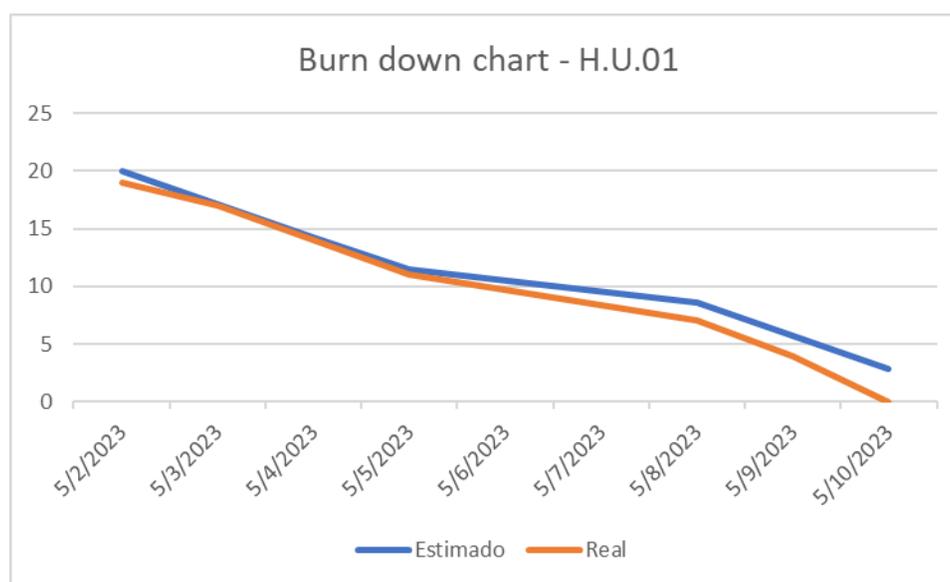
Burndown chart sprint 01

Ahora, para medir el avance del equipo, de forma cuantitativa, se hace uso de la herramienta Burndown Chart, o gráfico de agotamiento, esta es una herramienta esencial en la metodología Scrum que nos permite visualizar el progreso de un equipo durante un sprint específico (Rola & Kuchta, 2015). En el presente trabajo, este gráfico adquiere especial relevancia para garantizar que el equipo avance de manera eficiente y cumpla con los objetivos establecidos.

El Burndown Chart se usó para registrar el cumplimiento de las actividades diarias del Sprint 1, es decir, se usa para registrar el progreso de cada actividad desarrollada en el periodo comprendido del 2 al 10 de mayo de 2023, esto se representa en la Figura 18. Esta representación lineal nos ayuda a identificar, de manera oportuna, los posibles ajustes a realizar para los próximos Sprints.

Figura 18

Burndown chart Sprint 01



La Figura 18 presenta dos líneas distintas: la línea ideal y la línea real. La línea ideal representa la trayectoria que deberíamos seguir si quemáramos puntos de historia de manera uniforme a lo largo del sprint. Es decir, representa un progreso lineal, donde no están considerados posibles inconvenientes como: la dificultad técnica de implementar una tarea en específico, desorganización o falta de comunicación en el equipo, entre otros. Por otro lado, la línea real refleja el progreso auténtico del equipo, mostrando cuántos puntos de historia se han completado realmente cada día (Cohn, 2009).

Al analizar la Figura 18, es evidente que el equipo ha mantenido un ritmo cercano al ideal, con ligeras variaciones. Este patrón sugiere que hubo días donde se cumplió implementar tareas en el tiempo estimado y otros en los cuales no se logró, por ende, los días siguientes estuvieron cargados con

un extra de trabajo por realizar. Este tipo de información fue fundamental, porque nos permitió identificar áreas de mejora y ajustar el tiempo estimado para Sprints futuros.

Resultados Sprint 01.

Cada Sprint tiene como objetivo entregar un producto útil al usuario, en este caso el primer Sprint culminó en la obtención de los conjuntos de datos (datasets) necesarios para continuar con el desarrollo, tal y como se muestra en la Figura 19.

Figura 19

Videos capturados para la generación automática de datasets



Nota. Videos capturados para generar automáticamente los datasets empleados en el entrenamiento y validación de los modelos y algoritmos de machine learning utilizados para la Re-Identificación por características biométricas (rostro y forma de andar) y soft-biométricas (color, textura y silueta).

Sprint 2: Desarrollo de modelo facial

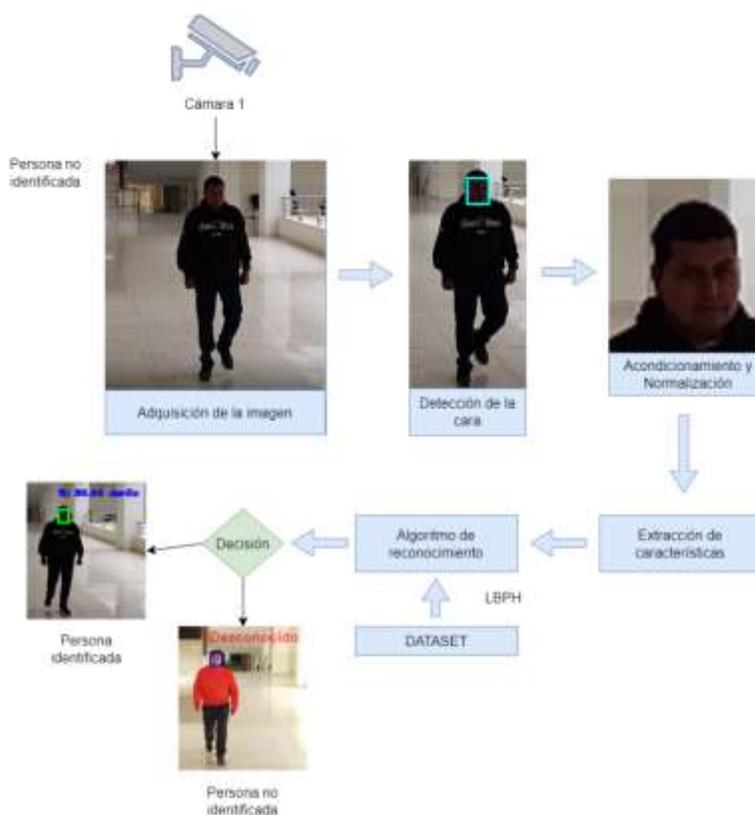
El Sprint 2, que tuvo lugar del 11 al 24 de mayo de 2023, se centró en el desarrollo del modelo facial, una pieza clave en el proceso de Re-Identificación de personas. La identificación facial es una

técnica biométrica ampliamente utiliza debido a la capacidad precisión que tienen estos sistemas por utilizar las características distintivas del rostro (Zhao *et al.*, 2003).

En la Figura 20 se muestra el esquema general de funcionamiento del sistema de sistema de identificación facial.

Figura 20

Descripción general del sistema de reconocimiento facial



Nota. El individuo es captado por la primera cámara, a continuación, se realiza la detección del rostro en las imágenes capturadas a lo largo del video para obtener el segmento de la imagen con el rostro de la persona, el cual representa la región de interés (ROI). Una vez delimitado el ROI, las imágenes resultantes son acondicionadas, normalizadas. Un conjunto de 300 imágenes es almacenadas para utilizarlos como dataset. Luego se realiza la extracción de características para utilizarlas como entrada

para el algoritmo de reconocimiento para luego realizar el entrenamiento del modelo para utilizarlo en la identificación del rostro.

Historia de usuario 02: Desarrollo de modelo facial.

En la historia de usuario H.U.02 se detalla que se requiere un modelo para reconocer personas a través del rostro, durante este Sprint se hace uso del conjunto de videos generados en el Sprint 1 para entrenar y validar el modelo que se deberá entregar al finalizar este Sprint.

Detalle Historia de Usuario 02. En este trabajo, cada Sprint se encuentra asociado a una determinada historia de usuario, a continuación, en la Tabla 9 se describen los elementos que componen la historia de usuario H.U.02, misma que cuenta con criterios de validación, responsables, estimación de tiempo y toda la información necesaria que permite identificar claramente el producto a entregar de este Sprint.

Tabla 9

Historia de usuario 02: Desarrollo de modelo facial

Historias de Usuario	
Numero: H.U.02	Usuario: Usuario
Nombre historia: Desarrollo de modelo facial	Numero Sprint: 2
Prioridad: Alta	Riesgo de desarrollo: Media
Puntos estimados (días): 10	interacción asignada: 1
Desarrolladores responsables: Bryan Chicaiza, Lenin Arteaga	
Descripción: Como usuario quiero un modelo y/o algoritmo de Visión por Computador que permita reconocer una persona mediante su rostro mientras se procesa la entrada de video de la primera cámara de seguridad.	
Validación:	

Historias de Usuario

- El modelo y/o algoritmo desarrollado reconoce el rostro de una persona.
 - Se evalúa el modelo y/o algoritmo desarrollado aplicando métricas evaluación.
 - Se prueba y ajusta el modelo para mejorar los resultados obtenidos.
-

Sprint Backlog 02

En el contexto de nuestra investigación sobre la Re-Identificación de personas, el Sprint 02, que se llevó a cabo del 11 al 24 de mayo de 2023, se centró en desarrollar y optimizar un modelo de reconocimiento facial que permita la Re-Identificación precisa de individuos a partir de imágenes y videos.

A continuación, en la Tabla 10 se detallan las tareas que se realizaron durante el Sprint 2, el cumplimiento de estas tareas en su totalidad permite la entrega del modelo que permite reconocer el rostro.

Tabla 10

Sprint Backlog 02

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.02	Elección de imágenes y videos faciales del dataset previamente segmentado, preprocesamiento	16	11/05/2023	12/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	de imágenes (estandarización de las imágenes)					
H.U.02	Implementación del algoritmo Haar Cascade/Viola- Jones para detectar rostros en las imágenes y videos.	16	15/05/2023	16/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.02	Entrenar el modelo de reconocimiento facial utilizando el conjunto de datos de entrenamiento.	24	17/05/2023	19/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.02	Validación y optimización del modelo (Ajuste de parámetros) de reconocimiento	24	22/05/2023	24/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

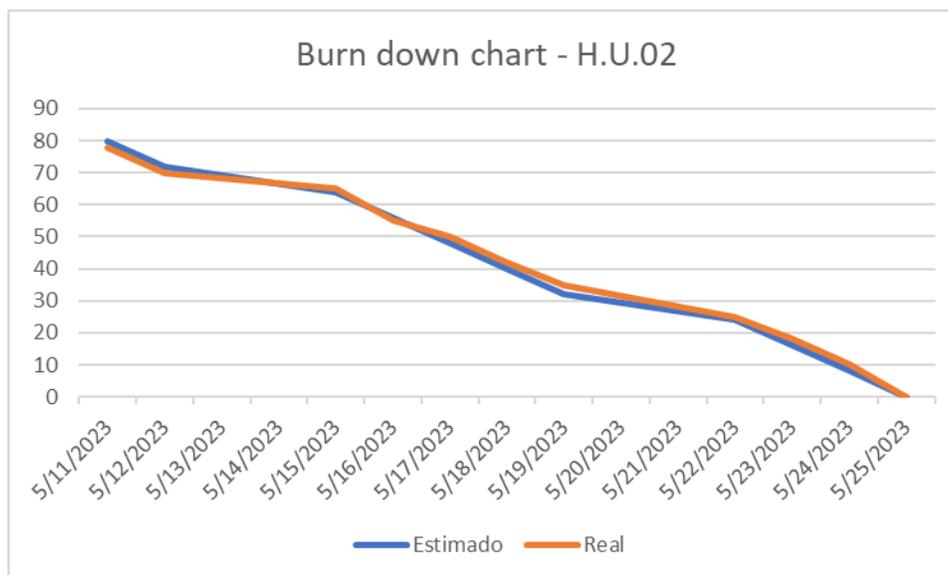
HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	facial, Monitoreo					
	de métricas de					
	evaluación.					

Burndown chart sprint 02

La Figura 21 del Sprint Burndown Chart para el Sprint 2 proporciona una representación visual del progreso del equipo durante el desarrollo del modelo facial. La línea ideal y la línea real representan, respectivamente, el progreso esperado y el progreso real del equipo a lo largo del sprint.

Figura 21

Burndown chart 02



Al analizar la Figura 21 se puede notar que la línea ideal comienza en 80 horas, que es el total de horas estimadas para este sprint, y disminuye uniformemente a lo largo de los 10 días laborables, reduciéndose en 8 horas por día. Esta línea representa el ritmo ideal de trabajo, donde el equipo completaría exactamente 8 horas de trabajo cada día.

Por otro lado, la línea real muestra el progreso real del equipo. En este escenario, el equipo comenzó cerca del ritmo ideal, pero hubo variaciones a lo largo del sprint. Por ejemplo, el 15 de mayo, el equipo solo pudo completar 5 horas de trabajo en lugar de las 8 horas ideales. Sin embargo, en días posteriores, como el 16 de mayo, se logró cumplir con las tareas al trabajar más de las 8 horas ideales.

Estas variaciones en la línea real significan que hubo pequeños retrasos en el cumplimiento de las tareas debido a inconvenientes encontrados durante el desarrollo de este Sprint, principalmente la dificultad técnica para ajustar el modelo facial y que proporcione predicciones acordes al estándar de la industria.

A pesar de las fluctuaciones en el progreso diario, la línea real converge con la línea ideal el 25 de mayo, lo que indica que el equipo logró completar todas las tareas planificadas para el sprint en el tiempo previsto.

Resultados Sprint 02

Cada Sprint tiene como objetivo entregar un producto útil al usuario, en este caso el segundo Sprint resultó en la obtención de un modelo de reconocimiento facial, no obstante, cada tarea comprendida en el Sprint Backlog 02 también permitió evidenciar el código generado para el componente de identificación facial Figura 23, el cual fue estructurado a través de su Pseudocódigo de la Figura 22, donde se representa la serie de instrucciones a seguir, para que el modelo de reconocimiento facial logre su cometido.

Figura 22

Pseudocódigo del algoritmo utilizado para la identificación del rostro

```

Entrada
Leer modelo entrenado:
Asignar ruta de la carpeta del dataset;
Leer video;
while:
  Leer fotograma del video;
  Redimensionar video;
  Convertir frame a escala de grises;
  Detectar rostro por frame con YOLO;
  For (alto, ancho, posición x mínima y y mínima) to parámetros do
    Definir área de interés;
    Obtener valor de confianza;
    If valor de confianza then:
      Dibujar recuadro del área de interés en frame;
      Imprimir nombre identificado;
    Else
      Dibujar recuadro del área de interés en frame;
      Imprimir sobre frame desconocido;
    End If
  End For
end while
Salida:
Mostrar frames del video;
Imprimir Verdaderos Positivos;
Imprimir Verdaderos Negativos;
Imprimir Falsos Positivos;
Imprimir Falsos Negativos;

```

Figura 23

Segmento de código utilizado para detectar el rostro y generar el dataset

```

for r in results:
    boxes = r.boxes
    for box in boxes:
        cls = int(box.cls[0])
        print("que imprime",cls)
        x1, y1, x2, y2 = box.xyxy[0]
        x1, y1, x2, y2 = int(x1), int(y1), int(x2), int(y2)
        rostro = auxFrame[y1:y2,x1:x2]

        cv2.rectangle(frame, (x1, y1), (x2, y2), (255, 255, 0), 2)

        rostro = cv2.resize(rostro,(150,150),interpolation=cv2.INTER_CUBIC)
        cv2.imwrite(personPath + '/rotro_{}.jpg'.format(count),rostro)
        count = count + 1
frame = imutils.resize(frame, width=1000)
cv2.imshow('YOLO V8 Detection', frame)
k=cv2.waitKey(1)
if k==27 or count>=2000:
    break

```

Luego de tener una idea general, representada en Pseudocódigo, se procedió a la elección y clasificación de las imágenes que fueron utilizadas para entrenar y probar el modelo, esto se ve en la Figura 24.

Figura 24

Imágenes generadas de la clase Danilo para el entrenamiento del algoritmo del rostro



Una vez obtenido el conjunto de datos necesario para alimentar el modelo, se desarrolló el código del modelo de reconocimiento facial, este comprende una serie de instrucciones en el lenguaje de programación Python, tal como se muestra en la Figura 25.

Figura 25

Segmento de código utilizado para el entrenar el algoritmo del rostro

```

labels = []
facesData = []
label = 0
for nameDir in peopleList:
    personPath = dataPath + '/' + nameDir
    print('Leyendo las imágenes')

    for fileName in os.listdir(personPath):
        print('Textura: ', nameDir + '/' + fileName)
        labels.append(label)
        facesData.append(cv2.imread(personPath+'/'+fileName,0)) #transformacion a escala de grises
    label = label + 1

X_train, X_test, y_train, y_test = train_test_split(facesData, np.array(labels),random_state=0)
# Métodos para entrenar el reconocedor
face_recognizer = cv2.face.LBPHFaceRecognizer_create()
# Entrenando el reconocedor de rostros
print("Entrenando...")
face_recognizer.train(X_train, np.array(y_train))

```

El resultado de la ejecución del código de programación para el modelo de reconocimiento facial es una estructura organizada de datos, esta estructura está compuesta por un conjunto de patrones los cuales son usados para detectar el rostro de una persona, tal y como se muestra en la Figura 26.

Figura 26

Modelo resultante del entrenamiento del algoritmo de identificación del rostro

```

Face.xml
1  <?xml version="1.0"?>
2  <opencv_storage>
3  <opencv_lbphfaces>
4    <threshold>1.7976931348623157e+308</threshold>
5    <radius>1</radius>
6    <neighbors>8</neighbors>
7    <grid_x>8</grid_x>
8    <grid_y>8</grid_y>
9    <histograms>

```

Al finalizar con el Sprint también se obtuvo el código fuente y ejecutable del componente solicitado (ver Figura 27 y Figura 28). Este código presenta una ventana que muestra la detección y porcentaje de reconocimiento del modelo para el reconocimiento facial

Figura 27

Código de identificación de un individuo en los videos que simulan las cámaras 1 y 2

```

for r in results:
boxes = r.boxes
for box in boxes:
cls = int(box.cls[0])
x1, y1, x2, y2 = box.xyxy[0]
x1, y1, x2, y2 = int(x1), int(y1), int(x2), int(y2)
rostro = auxFrame[y1:y2,x1:x2]
rostro = cv2.resize(rostro,(150,150),interpolation= cv2.INTER_CUBIC)#Rediseccionamiento de imagens de rostro
result = face_recognizer.predict(rostro)
print(result)

if result[1] < 60:#condición del valor de confianza
cv2.putText(frame,"%: {:.2f}".format(100-result[1])+" "+'{}'.format(imagePaths[result[0]]),(x1, y1-25),
cv2.rectangle(frame, (x1, y1), (x2, y2),(0,255,0,3)#cuando la condición es verdadera
print("RESULTADO: ",result[1])
else:
cv2.putText(frame,'Desconocido',(x1,y1-20),2,0.8,(0,0,255),1,cv2.LINE_AA)
cv2.rectangle(frame, (x1, y1), (x2, y2),(255, 0, 0),2)

```

Figura 28

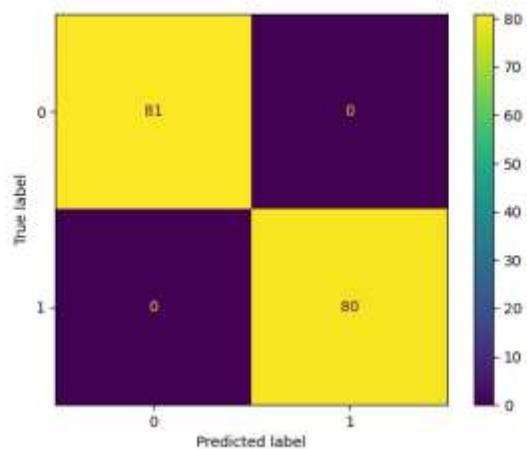
Identificación facial de una persona



Finalmente, se evaluó el modelo de reconocimiento facial mediante métricas de evaluación, esto nos permitió determinar que se cumplió a cabalidad los criterios de validación del Sprint 2, la matriz de confusión se muestra en la Figura 29 los resultados de las métricas de evaluación se aprecian en Tabla 11, estos valores son muy cercanos a 1 lo cual implica un alto nivel de reconocimiento del modelo de reconocimiento facial.

Figura 29

Matriz de confusión del resultado de identificación facial de la clase Danilo



Nota. Los resultados obtenidos en esta matriz de confusión corresponden al modelo de reconocimiento facial utilizando como entradas imágenes de la clase Danilo capturadas por la cámara 1. En la imagen se observa los siguientes resultados: 81 Verdaderos Positivos (0,0), 0 Falsos Negativos (0,1), 0 Falsos Positivos (1,0) y 80 Verdaderos Negativos (1,1). Aplicando la ecuación de la métrica accuracy (ver *Tabla 1*) se puede concluir que el porcentaje de predicción correspondiente a la clase Danilo es del 100 %. El resultado de las métricas accuracy, recall y precisión se pueden ver en la *Tabla 11*. Esta matriz es parte de los entregables de Sprint, la totalidad de las matrices se muestran en la sección Pruebas del algoritmo de identificación facial del capítulo IV.

Tabla 11

Resultados del algoritmo de identificación de personas por el rostro

Clasificador	Métricas	Valor
Haar	Accuracy	1.0000
Cascade	Recall	1.0000
	Precisión	1.000

Sprint 3: Desarrollo de modelo para la forma de caminar

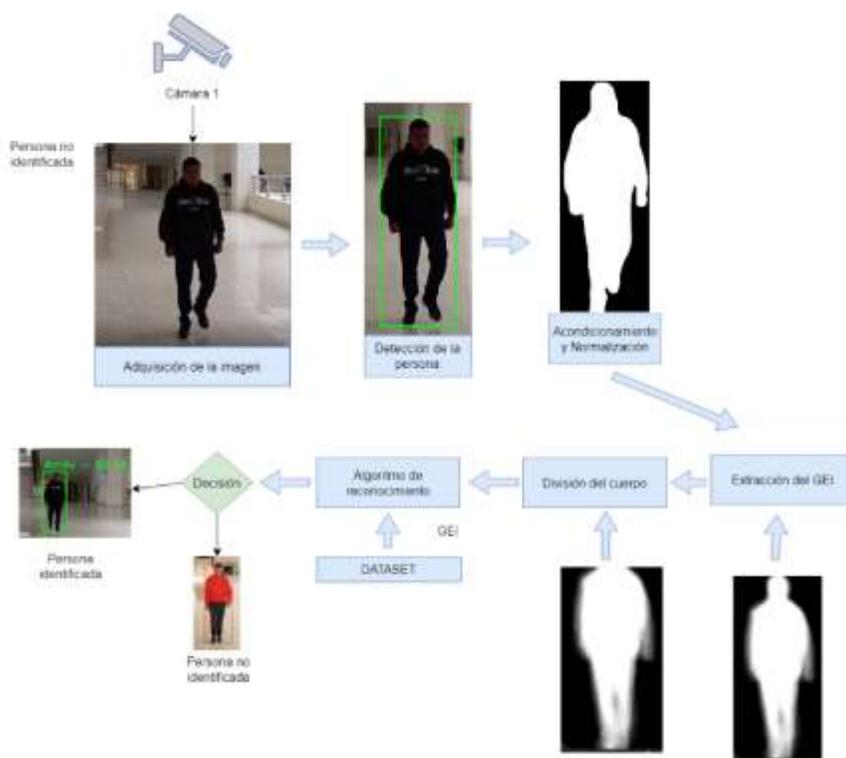
La forma en que una persona camina es única, es decir, es una característica biométrica esto permite determinar la identidad de una persona. En este sprint, el equipo se enfocó en capturar esta singularidad, desarrollando un modelo que pueda reconocer y diferenciar las formas de caminar de diferentes personas.

Historia de usuario 03: Desarrollo de modelo para la forma de caminar

La historia de usuario H.U.03 consiste en desarrollar un modelo que reconozca y distinga las formas de caminar de diferentes personas. A continuación, en la Figura 30 se muestra el esquema de funcionamiento del modelo de reconocimiento de personas según su forma de caminar, así también la Figura 31 muestra el esquema de Re-Identificación de personas según su forma de caminar.

Figura 30

Descripción general del sistema de Re-Identificación por la forma de andar (cámara 1)

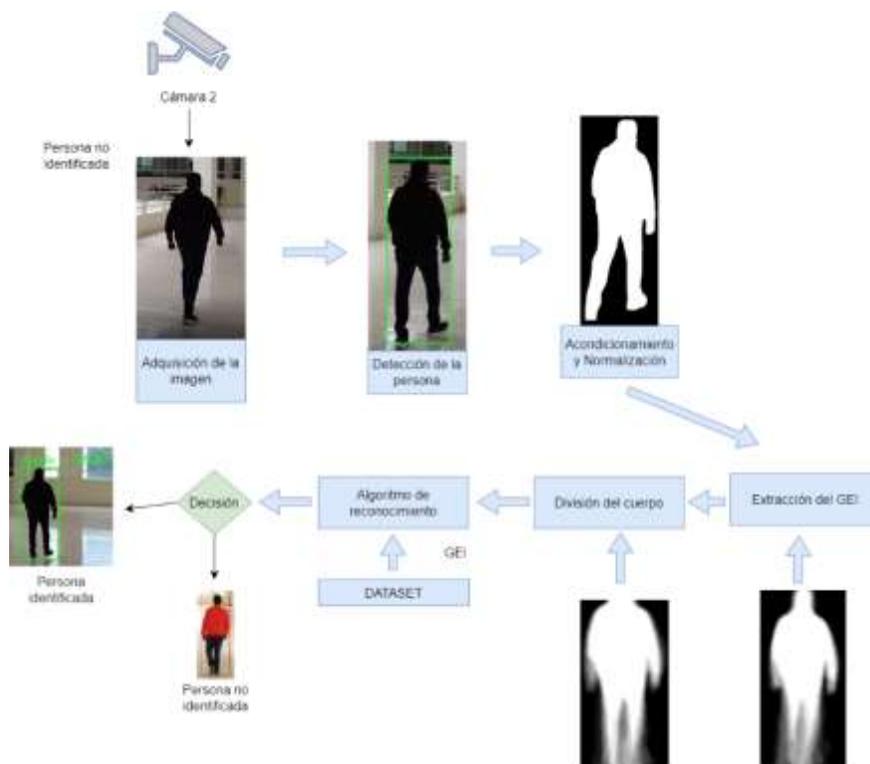


Nota. La persona es captada por la primera cámara, posteriormente, se realiza la detección de la persona en las imágenes capturadas a lo largo del video para obtener el segmento de la imagen de la persona (región de interés). Una vez delimitado el ROI, a las imágenes se realiza la extracción de la silueta para luego acondicionarlos y normalizados. Un conjunto de 600 imágenes es almacenado para utilizarlos como dataset. Como siguiente paso, se realiza la extracción del GEI y división del cuerpo de la

imagen para luego realizar el entrenamiento del modelo para utilizarlo en la Re-Identificación de personas.

Figura 31

Descripción general del sistema de Re-Identificación por la forma de andar (cámara 2)



Nota. En la segunda cámara, con las imágenes capturadas a lo largo del video se detecta a la persona para encontrar el área de la imagen donde se encuentra la persona (región de interés). Una vez delimitado el ROI, se extraen las siluetas de estas imágenes los cuales son acondicionas y normalizadas. Luego se realiza la extracción del GEI y división del cuerpo de la imagen. El algoritmo de reconocimiento cargará el modelo generado con las imágenes de dataset capturadas en la cámara 1. A continuación, el sistema realizará el reconocimiento con las imágenes del promedio de las siluetas obtenidas con las imágenes de la cámara 2.

Detalle Historia de Usuario 03. Esta historia de usuario consiste en desarrollar y optimizar un modelo que permita la Re-Identificación de individuos basado en su forma de caminar, utilizando técnicas de visión por computadora. La historia de usuario H.U.03 se muestra n la Tabla 12.

Tabla 12

Historia de usuario para el modelo de la forma de caminar

Historias de Usuario	
Numero: H.U.03	Usuario: Usuario
Nombre historia: Desarrollo de modelo para la forma de caminar	Numero Sprint: 3
Prioridad: Alta	Riesgo de desarrollo: Media
Puntos estimados (días): 12	interacción asignada: 1
Desarrolladores responsables: Bryan Chicaiza, Lenin Arteaga	
Descripción: Como usuario un modelo y/o algoritmo de Visión por Computador que permita reconocer una persona mediante su forma de caminar mientras se procesa la entrada de video de la primera cámara de seguridad.	
Validación:	
<ul style="list-style-type: none"> • Se identifica a una persona mediante su forma de caminar (característica biométrica). • Se evalúa el modelo y/o algoritmo desarrollado aplicando métricas evaluación. • Se prueba y ajusta el modelo para mejorar los resultados obtenidos. 	

Sprint Backlog 03

El Sprint 3, que se llevó a cabo del 25 de mayo de 2023 al 9 de junio de 2023, se centró en desarrollar y optimizar un modelo de reconocimiento por la forma de caminar que permita la Re-Identificación precisa de individuos a partir de imágenes y videos.

A continuación, en la Tabla 13 se detallan las tareas que se realizaron durante el Sprint 3, el cumplimiento de todas estas tareas resulta en la entrega del modelo que permita el reconocimiento de una persona mediante su forma de caminar.

Tabla 13

Sprint Backlog 03

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.03	Elección de imágenes y videos de la forma de caminar del dataset previamente segmentado, preprocesamiento de imágenes (estandarización de las imágenes)	16	25/05/2023	26/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.03	Implementación del algoritmo GEI para extraer características distintivas de la forma de caminar	24	29/05/2023	31/05/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

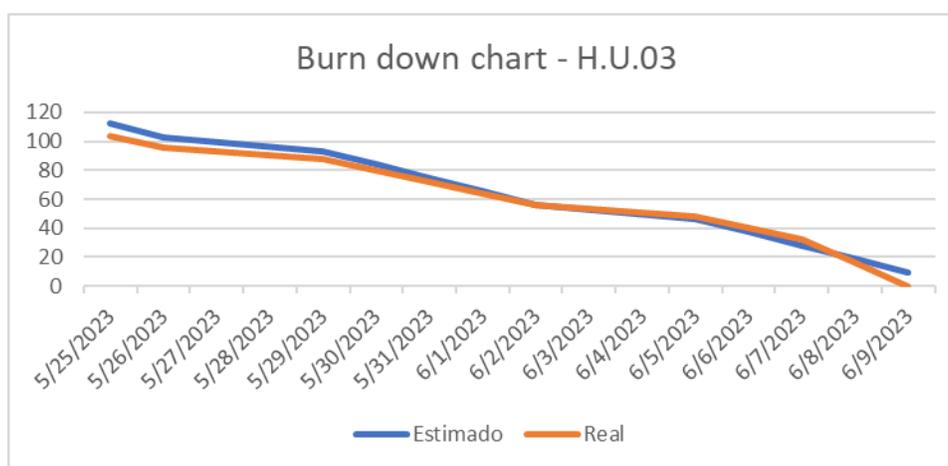
HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	en las imágenes y videos.					
H.U.03	Entrenar el modelo de reconocimiento por la forma de caminar utilizando el conjunto de datos de entrenamiento.	24	1/06/2023	5/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.03	Validación y optimización del modelo (Ajuste de parámetros) de reconocimiento por la forma de caminar, Monitoreo de métricas de evaluación.	32	6/06/2023	9/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

Burndown chart sprint 03

La Figura 32 del Sprint Burndown Chart para el Sprint 2 proporciona una representación visual del progreso del equipo durante el desarrollo del modelo de reconocimiento mediante la forma de caminar. La línea ideal y la línea real representan, respectivamente, el progreso esperado y el progreso real del equipo a lo largo del sprint.

Figura 32

Burndown chart sprint 03



En la Figura 32 la línea azul (la línea ideal) representa la trayectoria que debe seguir el equipo si se quiere realizar una determinada cantidad de trabajo cada día. A partir de la suma de las horas previstas para el sprint (112 horas), disminuye gradualmente hasta llegar a cero al final del periodo.

Por otro lado, el progreso real del equipo está representado por la línea naranja (línea real). En función del número de horas trabajadas, esta línea se ajusta cada día. Si esta línea es superior a la línea óptima, significa que no se está avanzando con lo previsto. Si es inferior, es posible que el equipo esté avanzando más rápido de lo previsto. En este caso se culminan las actividades en la fecha límite con un leve aumento de las tareas al finalizadas.

En el contexto del Sprint 3, se observa:

- El equipo comenzó con 112 horas de trabajo.

- A lo largo del sprint, hubo días en los que el progreso fue más lento que lo ideal, lo cual se debe a desafíos técnicos y ajustes en el modelo.

Resultados Sprint 03

Al finalizar este Sprint se obtuvieron los siguientes entregables: pseudocódigo del algoritmo utilizado para la Re-Identificación por la forma de andar (ver Figura 33), código fuente para la Re-Identificación por características por la forma de caminar (ver Figura 34), dataset para el entrenamiento (ver Figura 35), código para el entrenamiento (ver

Figura 36), modelo generado luego del entrenamiento (ver Figura 37), matriz con los resultados de las métricas: exactitud (accuracy), recall y precisión del modelo (ver Tabla 15 y Figura 38, Figura 50), y la nueva versión del sistema con el componente de Re-Identificación de personas por la forma de caminar (ver Figura 39 y Figura 40).

Figura 33

Pseudocódigo del algoritmo utilizado para la identificación de la forma de andar

```

Entrada
Leer modelo entrenado;
Asignar ruta de la carpeta del dataset;
Leer video;
while:
  Leer fotograma del video;
  Detectar objetos por frame con YOLO;
  For (alto, ancho, posición x mínima y y mínima) to parámetros do
    Definir área de interés;
    Obtener siluetas;
    If se encuentra en 50 frames:
      Reiniciar a 0 contador;
      Obtener promedio de las siluetas;
      Obtener índice mayor precisión;
      Obtener nombre de persona identificada;
      If valor de confianza es mayor a 80 por ciento then:
        Dibujar recuadro del área de interés en frame;
        Imprimir nombre identificado;
      Else
        Dibujar recuadro del área de interés en frame;
        Imprimir sobre frame desconocido;
      End If
    End If
  End For
end while
Salida:
Mostrar frames del video;
Imprimir Verdaderos Positivos;
Imprimir Verdaderos Negativos;
Imprimir Falsos Positivos;
Imprimir Falsos Negativos;

```

Figura 34

Código utilizado para detectar una persona, obtener su silueta y generar el dataset

```

# filter the detections to be only people
if class_names[class_id] != "person":
    continue

xmin, ymin, xmax, ymax = int(data[0]), int(data[1]), int(data[2]), int(data[3])
person = frame[ymin:ymax, xmin:xmax]
#silhouette = preprocessing.background_subtraction(firstFrame,frame, (xmin, ymin, xmax, ymax))
silhouette = preprocessing.get_silhouette(person, (xmin, ymin, xmax, ymax))
#cv2.imshow("silhouette", silhouette)
if(count==55):
    count = 0
    gei_redim = preprocessing.get_gei(images=silhouette_data)
    val = tf.stack([gei_redim])
    prediction = model_sa.predict(val)
    prediction = np.array(prediction[0])

```

Figura 35

Imágenes generadas para el entrenamiento del modelo de la forma de andar

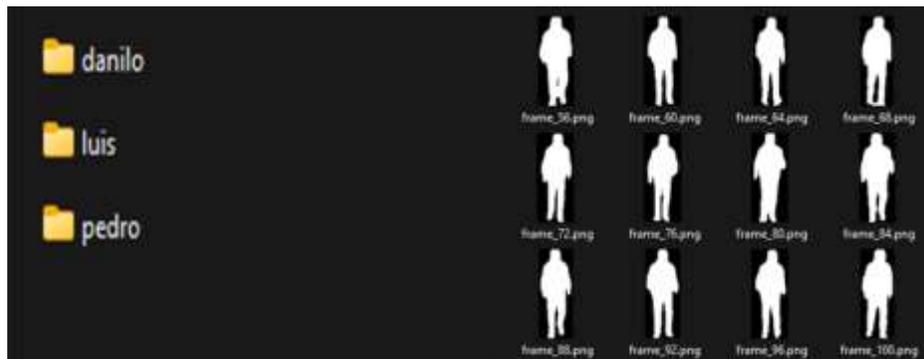


Figura 36

Segmento de código utilizado para entrenar el modelo de la forma de andar

```

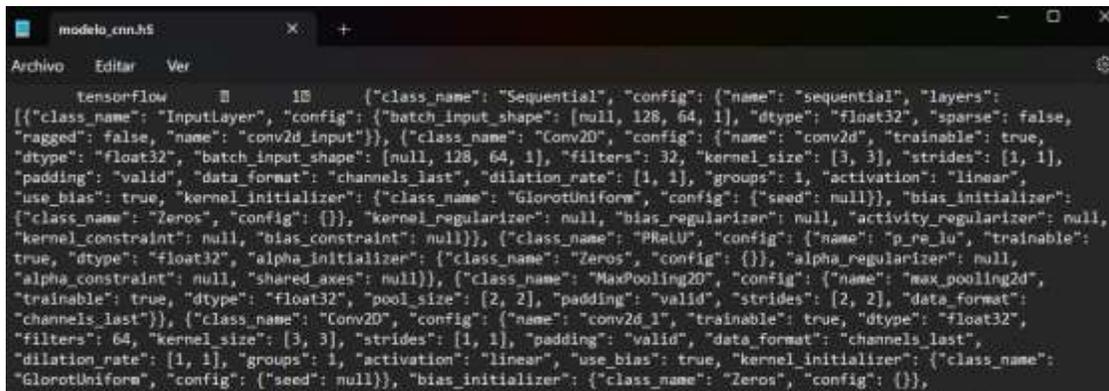
directories = ['nm-01', 'nm-02', 'nm-03', 'nm-04']
video_angles = ['000']
for angle in video_angles:
    classes = os.listdir("Silhouette/"+str(angle)+"/")
    for i, current_class in enumerate(classes):
        for directory in directories:
            path = "Silhouette/"+str(angle)+"/"+current_class+"/"+directory+"/"
            #print(path)
            files = os.listdir(path)
            images = [imread(path+f) for f in files]
            gei_redim = Preprocessing().get_gei(images=images)
            data.append(gei_redim)
            labels.append(i)

import tensorflow as tf
gait_model = GaitModel()
X_train, y_train, X_test, y_test, X_val, y_val = gait_model.split_dataset(data=data, labels = labels)
print(y_train)
print(y_test)
print(y_val)
history = gait_model.train(data=data, labels = labels)
loss, accuracy = gait_model.get_model().evaluate(X_test, y_test, verbose=1)

```

Figura 37

Modelo resultante de Re-Identificación por la forma de andar.



```

tensorflow 1.15.1 { "class_name": "Sequential", "config": { "name": "sequential1", "layers":
[{"class_name": "InputLayer", "config": {"batch_input_shape": [null, 128, 64, 1], "dtype": "float32", "sparse": false,
"ragged": false, "name": "conv2d_input"}}, {"class_name": "Conv2D", "config": {"name": "conv2d", "trainable": true,
"dtype": "float32", "batch_input_shape": [null, 128, 64, 1], "filters": 32, "kernel_size": [3, 3], "strides": [1, 1],
"padding": "valid", "data_format": "channels_last", "dilation_rate": [1, 1], "groups": 1, "activation": "linear",
"use_bias": true, "kernel_initializer": {"class_name": "GlorotUniform", "config": {"seed": null}}, "bias_initializer":
{"class_name": "Zeros", "config": {}}, "kernel_regularizer": null, "bias_regularizer": null, "activity_regularizer": null,
"kernel_constraint": null, "bias_constraint": null}, {"class_name": "PReLU", "config": {"name": "p_re_lu", "trainable":
true, "dtype": "float32", "alpha_initializer": {"class_name": "Zeros", "config": {}}, "alpha_regularizer": null,
"alpha_constraint": null, "shared_axes": null}, {"class_name": "MaxPooling2D", "config": {"name": "max_pooling2d",
"trainable": true, "dtype": "float32", "pool_size": [2, 2], "padding": "valid", "strides": [2, 2], "data_format":
"channels_last"}}, {"class_name": "Conv2D", "config": {"name": "conv2d_1", "trainable": true, "dtype": "float32",
"filters": 64, "kernel_size": [3, 3], "strides": [1, 1], "padding": "valid", "data_format": "channels_last",
"dilation_rate": [1, 1], "groups": 1, "activation": "linear", "use_bias": true, "kernel_initializer": {"class_name":
"GlorotUniform", "config": {"seed": null}}, "bias_initializer": {"class_name": "Zeros", "config": {}},

```

Tabla 14

Código utilizado para la Re-Identificación de un individuo por la forma de andar

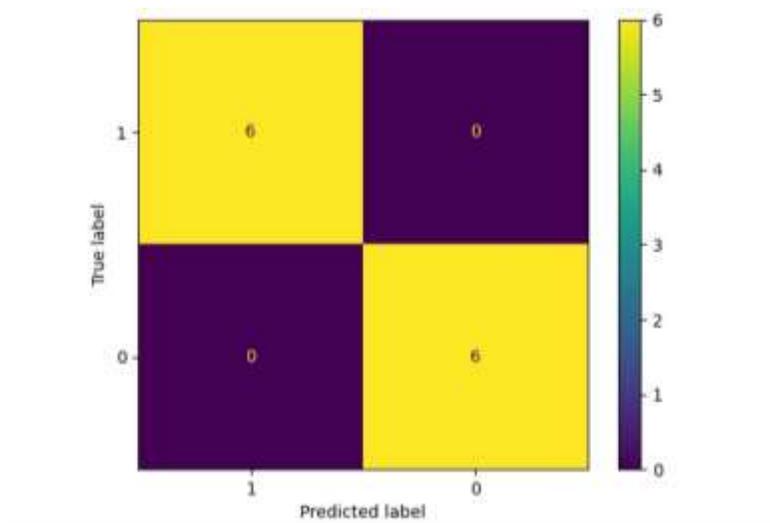
```

if(count==55):
    count = 0
    gei_redim = preprocessing.get_gei(images=sihoulette_data)
    val = tf.stack([gei_redim])
    prediction = model_sa.predict(val)
    prediction = np.array(prediction[0])
    try:
        presition = max(prediction)
        index = np.where(prediction==presition)[0][0]
        label = labels[index]+ ' - ' + str(round(max(prediction)*100, 2))
        if(round(max(prediction)*100, 2)<80):
            label = ""
    except:
        pass
    sihoulette_data = []

```

Figura 38

Matriz de confusión del resultado de Re-Identificación por la forma de andar de la clase Danilo



Nota. Los resultados obtenidos en esta matriz de confusión corresponden al modelo de reconocimiento facial utilizando como entradas imágenes de la clase Danilo capturadas por la cámara 1. En la imagen se observa los siguientes resultados: 6 Verdaderos Positivos (0,0), 0 Falsos Negativos (0,1), 0 Falsos Positivos (1,0) y 6 Verdaderos Negativos (1,1). Aplicando la ecuación de la métrica accuracy (ver *Tabla 1*) se puede concluir que el porcentaje de predicción correspondiente a la clase Danilo es del 100%, utilizando las imágenes de la cámara 1. El resultado de las métricas accuracy, recall y precisión se pueden ver en la *Tabla 15*. Esta matriz es parte de los entregables de este Sprint, la totalidad de las matrices se encuentran en la sección Pruebas del algoritmo de Re-Identificación forma de andar del capítulo IV.

Tabla 15

Resultados del algoritmo de Re-Identificación por la forma de andar de la clase Danilo

Clasificador	Métricas	Valor
CNN	Accuracy	0.9520
	Recall	0.9440
	Precisión	0.9440

Figura 39

Identificación facial de una persona por la forma de andar, cámara 1

**Figura 40**

Identificación facial de una persona por la forma de andar, cámara 2



Sprint 4: Desarrollo de Modelos Soft-biométricos

En el cuarto sprint del sistema de Re-Identificación de personas, nos adentramos en un aspecto crucial para la Re-Identificación de personas: las características soft-biométricas. Estas características,

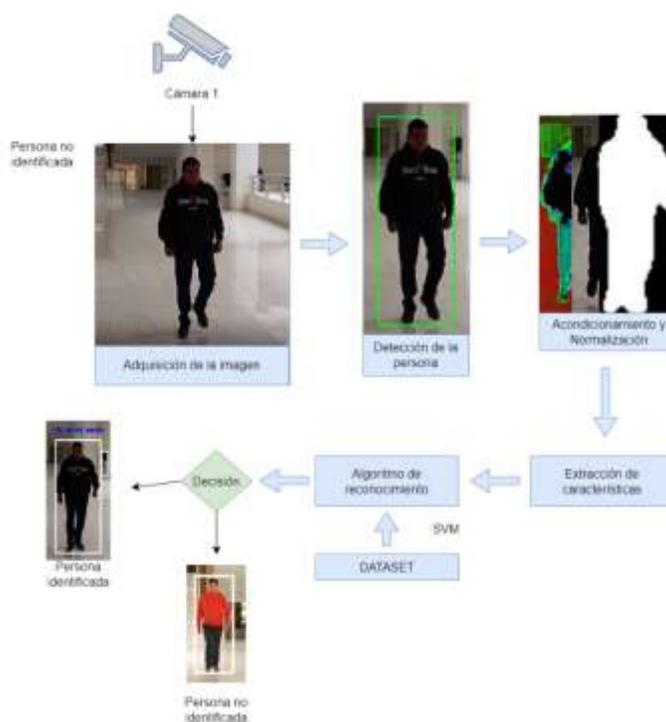
que incluyen elementos como la silueta corporal y la textura y color de la vestimenta, ofrecen una dimensión adicional de datos que puede ser esencial para identificar a individuos en situaciones donde las características biométricas tradicionales, como el rostro o la forma de caminar, no son suficientemente distintivas o están parcialmente ocultas.

Al integrar estos modelos Soft-biométricos en nuestro sistema, buscamos no solo mejorar la precisión general de la Re-Identificación, sino también hacer que nuestro sistema sea más robusto y adaptable a una variedad de escenarios y condiciones. En este sprint, nos centraremos en desarrollar, validar e integrar estos modelos, asegurando que complementen y potencien las capacidades de los modelos previamente desarrollados.

En la Figura 41 y Figura 42 se muestra el esquema general que utilizado para el desarrollo del componente que permita la Re-Identificación por las características Soft-biométricas.

Figura 41

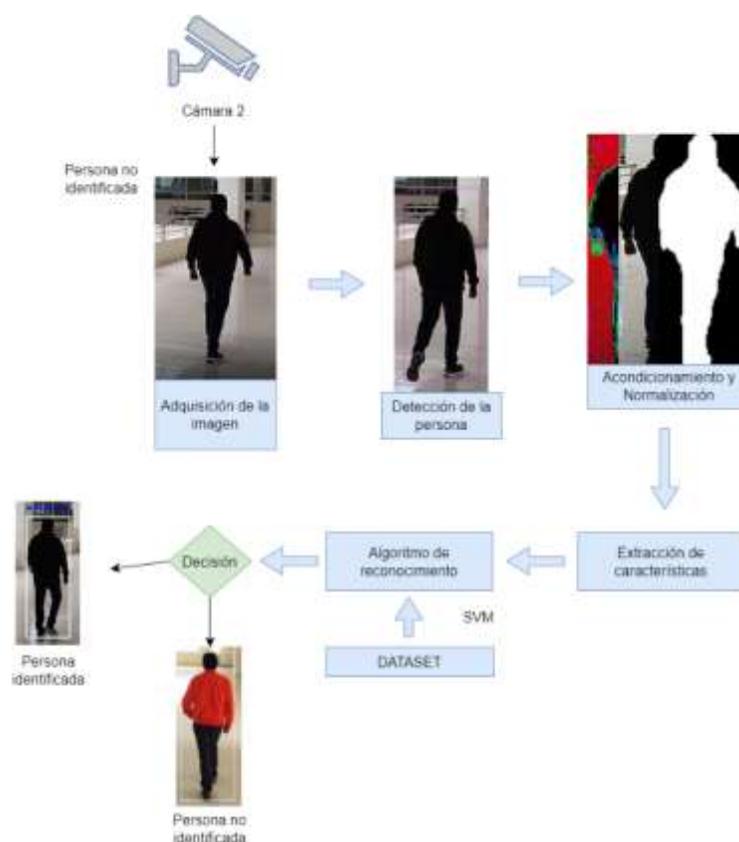
Esquema de Re-Identificación por características Soft-biométricas (cámara 1)



Nota. La persona es captada por la primera cámara, a continuación, se realiza la detección de la persona en las imágenes capturadas a lo largo del video para obtener el segmento de la imagen de la persona (región de interés). Una vez delimitado el ROI, las imágenes resultantes son acondicionadas, normalizadas, para posteriormente obtener la silueta y la imagen en modelo HSV. Un conjunto de 600 imágenes (300 por cada característica Soft-biométrica) son almacenado para utilizarlos como dataset para luego realizar el entrenamiento del modelo para utilizarlo en la Re-Identificación de personas.

Figura 42

Descripción del sistema de Re-Identificación por características Soft-biométricas (cámara 2)



Nota. En la segunda cámara, con las imágenes capturadas a lo largo del video se detecta al individuo para encontrar el área de la imagen donde se encuentra la persona (región de interés). Una vez delimitado el ROI, se realiza el acondicionamiento y normalización de la imagen resultante para posteriormente obtener la silueta y la imagen en modelo HSV. El algoritmo de reconocimiento cargará el

modelo generado con las imágenes de dataset capturadas en la cámara 1. A continuación, se realiza la Re-Identificación con las características soft-biométricas (color, textura y silueta) extraídas de las imágenes capturadas por la cámara 2.

Historia de Usuario 04: Desarrollo de Modelos Soft-biométricos

La historia de usuario H.U.04 consiste en desarrollar modelos que se centren en características soft-biométricas, como la silueta corporal, el color y la textura de la vestimenta, para mejorar la precisión y robustez del sistema de Re-Identificación.

Detalle Historia de Usuario 04. La historia de usuario que está involucrada en este Sprint es la H.U.04, la cual se encuentra en la Tabla 16.

Tabla 16

Historia de usuario para el modelo Soft-biométrico

Historias de Usuario	
Numero: H.U.04	Usuario: Usuario
Nombre historia: Desarrollo de modelo soft-biométrico	Numero Sprint: 4
Prioridad: Alta	Riesgo de desarrollo: Media
Puntos estimados (días): 13	interacción asignada: 1
Desarrolladores responsables: Bryan Chicaiza, Lenin Arteaga	
Descripción: Como usuario quiero un modelo y/o algoritmo de Visión por Computador que permita reconocer una persona mediante su forma de caminar mientras se procesa la entrada de video de la primera cámara de seguridad.	
Validación:	
<ul style="list-style-type: none"> • Se identifica a una persona mediante su silueta corporal. 	

Historias de Usuario

- Se identifica a una persona mediante el color de la vestimenta.
 - Se identifica a una persona mediante la textura de la vestimenta.
 - Se evalúa el modelo y/o algoritmo desarrollado aplicando métricas evaluación.
 - Se prueba y ajusta el modelo para mejorar los resultados obtenidos.
-

Sprint Backlog 04

El Sprint 4, que se llevó a cabo del 12 de junio de 2023 al 28 de junio de 2023, se centró en desarrollar y optimizar un modelo de reconocimiento por características soft-biométricas (silueta corporal, color y textura de la vestimenta) que permita la Re-Identificación precisa de individuos a partir de imágenes y videos.

La Tabla 17 las tareas que se realizaron durante el Sprint 4.

Tabla 17

Sprint Backlog 04

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.04	Elección de imágenes y videos de características soft-biométricas del dataset previamente segmentado, preprocesamiento de imágenes	16	12/06/2023	13/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	(estandarización de las imágenes)					
H.U.04	Implementación de los algoritmos LBP, HOG, SVM, CNN, YOLO para extraer características distintivas de la silueta corporal, color y textura de la vestimenta en las imágenes y videos.	56	14/06/2023	22/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.04	Entrenar el modelo de reconocimiento por características soft-biométricas (silueta corporal, color y textura de la vestimenta) utilizando el	16	23/06/2023	26/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	conjunto de datos de entrenamiento.					
H.U.04	Validación y optimización del modelo (Ajuste de parámetros) de reconocimiento por características Soft-biométricas (silueta corporal, color y textura de la vestimenta), Monitoreo de métricas de evaluación.	16	27/06/2023	28/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

Burndown chart sprint 04

La Figura 43 proporciona una representación visual del trabajo pendiente en el sprint en relación con el tiempo. En este gráfico, se comparan dos líneas: la línea "Estimado (Ideal)" y la línea "Real".

Figura 43

Burndown chart sprint 04.



Al analizar la Figura 43 se observa que, al comparar los resultados obtenidos entre las dos líneas, el equipo ha tenido un rendimiento bastante cercano al ideal, aunque con algunas variaciones. Se observa que, al inicio del sprint, el equipo trabajó a un ritmo más rápido que el ideal, pero hacia el final, se alineó más estrechamente con la línea ideal, terminando todas las tareas sin mayor dificultad el día límite establecido.

Resultados Sprint 04

Al finalizar este sprint se obtuvieron los siguientes entregables: pseudocódigo del algoritmo utilizado para la identificación Soft-biométrica (ver Figura 44), código fuente para la Re-Identificación por características Soft-biométricas (ver Figura 45), dataset para el entrenamiento (ver Figura 46), código para el entrenamiento (ver Figura 47), modelo generado luego del entrenamiento (ver Figura 48), matriz con los resultados de las métricas: exactitud (accuracy), recall y precisión del modelo (ver Tabla 18 y Figura 50), y la nueva versión del sistema con el componente de Re-Identificación de personas por características Soft-biométricas (ver Figura 51 y Figura 52).

Figura 44

Pseudocódigo del algoritmo para la identificación Soft-biométrica

```

Entrada
Leer modelo entrenado;
Asignar ruta de la carpeta del dataset;
Leer video;
while:
  Leer fotograma del video;
  Convertir frame a escala de grises;
  Detectar objetos por frame con YOLO;
  Convertir BGR a HSV;
  For (alto, ancho, posición x mínima y y mínima) to parámetros de
    Definir área de interés de textura;
  Definir área de interés de color;
  Definir área de interés de silueta;
  Obtener mascara binaria del área de interés;
  Redimensionar mascara binaria;
  Obtener valor de confianza de textura;
  Obtener valor de confianza de color;
  Obtener valor de confianza de silueta;
  Obtener mejor valor de confianza;
  If mejor valor de confianza then:
    Dibujar recuadro del área de interés en frame;
    Imprimir nombre identificado;
  Else
    Dibujar recuadro del área de interés en frame;
    Imprimir sobre frame desconocido;
  End If
  End For
end while
Salida:
Mostrar frames del video;
Imprimir Verdaderos Positivos;
Imprimir Verdaderos Negativos;
Imprimir Falsos Positivos;
Imprimir Falsos Negativos;

```

Figura 45

Código utilizado para obtener silueta corporal, el color y textura de la vestimenta para generar el dataset

```

video="Videos/danilo.mp4"
name="Alex"
def Color(): ...

def Textura(): ...

def Silueta(): ...

def runInParallel(*fns):
    proc = []
    for fn in fns:
        p = Process(target=fn)
        print("PROCESADORES",p)
        p.start()
        proc.append(p)
    for p in proc:
        print("SALIDA PROCESADORES",p)
        p.join()

if __name__ == '__main__':
    runInParallel(Textura,Color,Silueta)

```

Figura 46

Imágenes generadas para el entrenamiento del modelo de características Soft-biométricas



Figura 47

Segmento de código utilizado para entrenar el modelo soft-biométrico combinado

```
def train():
    dataPath = 'Data_CTS' #Cambia a la ruta donde hayas almacenado Data
    peopleList = os.listdir(dataPath)
    print('Lista de personas: ', peopleList)

    labels = []
    facesData = []
    label = 0
    for nameDir in peopleList:...

    # Métodos para entrenar el reconocedor
    face_recognizer = cv2.face.LBPHFaceRecognizer_create()

    # Entrenando el reconocedor de rostros
    print("Entrenando...")
    face_recognizer.train(facesData, np.array(labels))

    # Almacenando el modelo obtenido
    face_recognizer.write('multiple.xml')
    print("Modelo almacenado...")
```

Figura 48

Modelo resultante por características Soft-biométricas

```

multiple.xml
1  <?xml version="1.0"?>
2  <opencv_storage>
3  <opencv_lbphfaces>
4  <threshold>1.7976931348623157e+308</threshold>
5  <radius>1</radius>
6  <neighbors>8</neighbors>
7  <grid_x>8</grid_x>
8  <grid_y>8</grid_y>
9  <histograms>
10 <_ type_id="opencv-matrix">
11 <rows>1</rows>
12 <cols>16384</cols>
13 <dt>f</dt>
14 <data>
15     2.85714287e-02 0. 0. 0. 5.71428575e-02 0. 0. 7.14285718e-03 0.

```

Figura 49

Código utilizado para la Re-Identificación de un individuo por características Soft-biométricas aras 1 y 2

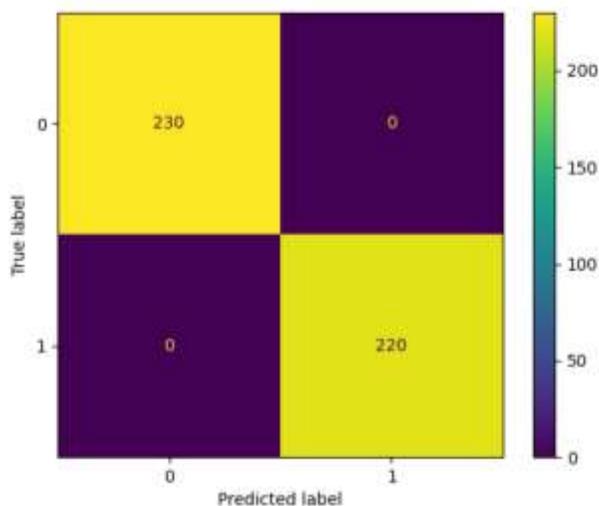
```

conf_silueta = 1 - result_silueta[1] / 100
conf_textura = 1 - result_textura[1] / 100
conf_color = 1 - result_color[1] / 100
min_conf = min(conf_silueta, conf_textura, conf_color)
if min_conf == conf_silueta:
    index = result_silueta[0]
elif min_conf == conf_textura:
    index = result_textura[0]
else:
    index = result_color[0]
# Obtener la etiqueta de la imagen correspondiente
etiqueta = imagePaths[index]
print("resultado_silueta", result_silueta)
print("resultado_textura", result_textura)
print("resultado_color", result_color)
results = [(result_silueta[0], result_silueta[1]), (result_textura[0], result_textura[1]), (result_color[0], result_color[1])]
results = sorted(results, key=lambda x: x[1], reverse=True)
print("resultado del combinado", results)
if results[0][1] < 70:
    cv2.putText(img, "X: {:.2f}".format(100 - results[0][1]) + " + {}".format(imagePaths[results[0][0]]), (x1, y1))
    if imagePaths[results[0][0]].startswith('desconocido'):
        print("Persona desconocida")
    else:
        print("Persona conocida")

```

Figura 50

Matriz de confusión resultante para la clase Danilo



Nota. Los resultados obtenidos en esta matriz de confusión corresponden al modelo de reconocimiento facial utilizando como entradas imágenes de la clase Danilo capturadas por la cámara 1. En la imagen se observa los siguientes resultados: 230 Verdaderos Positivos (0,0), 0 Falsos Negativos (0,1), 0 Falsos Positivos (1,0) y 220 Verdaderos Negativos (1,1). Aplicando la ecuación de la métrica accuracy (ver *Tabla 1*) se puede concluir que el porcentaje de predicción correspondiente a la clase Danilo es del 100%, utilizando las imágenes de la cámara 1. El resultado de las métricas accuracy, recall y precisión se pueden ver en la *Tabla 15*. Esta matriz es parte de los entregables de este Sprint, la totalidad de las matrices se encuentran en la sección Pruebas del algoritmo de Re-Identificación de color, textura y silueta del capítulo IV

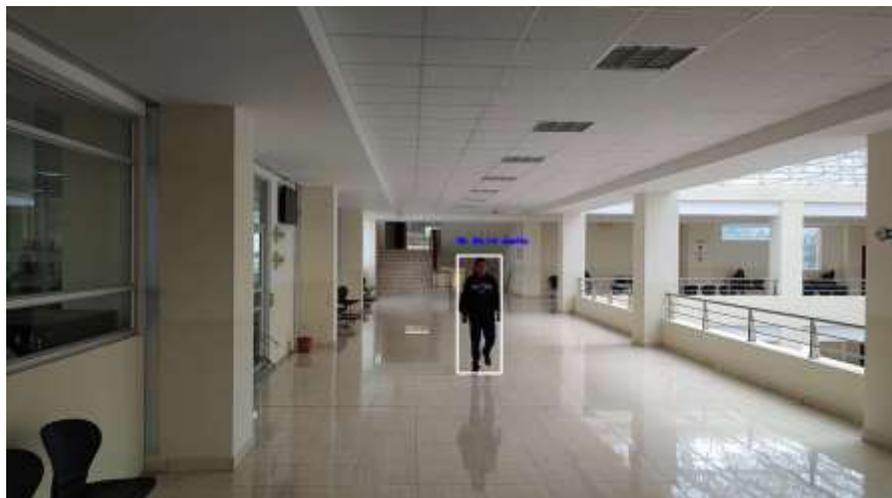
Tabla 18

Resultados del algoritmo de Re-Identificación por características soft-biométricas

Clasificador	Métricas	Valor
SVM	Accuracy	1.0000
	Recall	1.0000
	Precisión	1.0000

Figura 51

Identificación de una persona por características soft-biométricas (color, textura, silueta), cámara 1

**Figura 52**

Identificación de una persona por características soft-biométricas (color, textura, silueta), cámara 2



Sprint 5: Fusión de Modelos

El Sprint 5 se centra en combinar el poder y la precisión de cada modelo individual para crear un sistema integrado y robusto. Esta fase es esencial, ya que la combinación adecuada de características faciales, de forma de caminar y soft-biométricas puede potenciar enormemente la precisión del sistema. Para ello, se debe asegurar que la fusión no solo sea efectiva sino también eficiente.

Historia de Usuario 05: Fusión de Modelos

La historia de usuario H.U.05 consiste en fusionar los modelos desarrollados previamente, es decir, el modelo facial, el modelo de forma de caminar y los modelos soft-biométricos, para crear un sistema integrado de Re-Identificación de personas. Esta fusión permitirá que el sistema aproveche la información combinada de estas características distintas para mejorar la precisión y robustez en la Re-Identificación.

Detalle Historia de Usuario 05. *La historia de usuario que está involucrada en este Sprint es la H.U.05, la cual se encuentra en la Tabla 19.*

Tabla 19

Historia de usuario para fusión de modelos

Historias de Usuario	
Numero: H.U.05	Usuario: Usuario
Nombre historia: Desarrollo de modelo soft-biométrico	Numero Sprint: 5
Prioridad: Alta	Riesgo de desarrollo: Media
Puntos estimados (días): 10	interacción asignada: 1
Desarrolladores responsables: Bryan Chicaiza, Lenin Arteaga	
Descripción: Como usuario quiero un modelo y/o algoritmo de Visión por Computadora que permita identificar una persona mediante características biométricas (rostro y forma de caminar) y soft-biométricas (silueta corporal, color y textura de la vestimenta) mediante la fusión de modelos y/o algoritmos de Re-Identificación.	
Validación:	
<ul style="list-style-type: none"> • Se identifica a una persona mediante el modelo biométrico (rostro y forma de caminar). 	

Historias de Usuario

- Se identifica a una persona mediante el modelo soft-biométrico (silueta corporal, color y textura de la vestimenta).
 - Se evalúa el modelo y/o algoritmo desarrollado aplicando métricas evaluación.
 - Se prueba y ajusta el modelo para mejorar los resultados obtenidos.
-

Sprint Backlog 05

El Sprint 5, que se llevó a cabo del 29 de junio de 2023 al 12 de julio de 2023, se centró en desarrollar y optimizar un modelo de reconocimiento por características biométricas (rostro y forma de caminar) y soft-biométricas (silueta corporal, color y textura de la vestimenta) que permita la Re-Identificación precisa de individuos a partir de imágenes y videos, a través de la fusión de los modelos previamente desarrollados en los sprints anteriores.

Las tareas que se realizaron durante el Sprint 5 se presentan en **Tabla 20**.

Tabla 20

Sprint Backlog 05

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.05	Revisión y preparación de modelos biométricos y soft-biométricos, verificar optimización.	16	29/06/2023	30/06/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

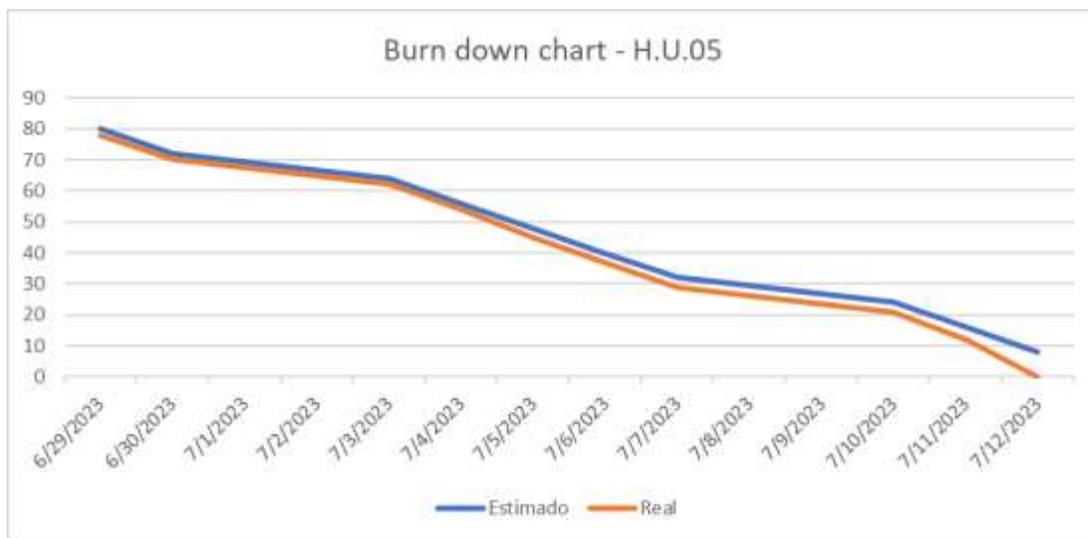
HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.05	Integración de modelos en una estructura de predicción común, verificar interacción.	24	3/07/2023	5/07/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.05	Ajustar los pesos y parámetros para maximizar la precisión de la Re-Identificación.	16	6/07/2023	7/07/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.05	Validación y optimización del modelo fusionado, Identificar incompatibilidades en la lógica de fusión, Monitoreo de métricas de evaluación.	24	10/07/2023	12/07/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

Burndown chart sprint 05

La Figura 53 proporciona una representación visual del progreso del equipo durante el periodo del 29 de junio al 12 de julio de 2023. En el eje vertical (Y) del gráfico, se representan las horas de trabajo pendientes, comenzando con un total de 80 horas, que es la estimación inicial de trabajo para este sprint. El eje horizontal (X) representa los días laborables del sprint, desde el inicio hasta el final.

Figura 53

Burndown chart sprint 05



En la Figura 53, podemos observar que para este sprint la línea real sigue de cerca a la línea ideal, con algunas variaciones. Estas fluctuaciones en la línea real se deben a diversos factores, como desafíos técnicos inesperados, en relación con el acoplamiento de los modelos biométricos y soft-biométricos.

Resultados Sprint 05

Al finalizar este sprint se obtuvieron los siguientes entregables: pseudocódigo del algoritmo utilizado para fusionar los tres modelos generados (ver

Figura 54), algoritmo para la Re-Identificación de personas con la fusión de los modelos y/o algoritmos de Machine Learning (ver Figura 55) y la nueva versión del sistema con el componente planificado para este sprint (ver Figura 56 y Figura 57).

Figura 54

Pseudocódigo del algoritmo para fusionar los modelos a través del promedio ponderado

```

Entrada
Asignar ruta de la carpeta del dataset;
Leer video;
while:
    Leer fotograma del video;
    Convertir frame a escala de grises;
    Detectar objetos por frame con YOLO;
    Convertir BGR a HSV;
    For (alto, ancho, posición x mínima y y mínima) to parámetros de
        Definir área de interés de textura;
Definir área de interés de color;
        Definir área de interés de persona;
        Obtener siluetas;
        Obtener promedio de las siluetas;
        Obtener predicción de la fusión de los 3 modelos;
        If valor de predicción es mayor a 70 then:
            Dibujar recuadro del área de interés en frame;
            Imprimir nombre identificado;
        Else
            Dibujar recuadro del área de interés en frame;
            Imprimir sobre frame desconocido;
        End If
    End For
end while
Salida:
    Mostrar frames del video;
    Imprimir Verdaderos Positivos;
    Imprimir Verdaderos Negativos;
    Imprimir Falsos Positivos;
    Imprimir Falsos Negativos;

```

Figura 55

Código utilizado para realizar la fusión de las características a través del promedio ponderado

```

# Modelo combinado Color, Silueta, Textura
multiple_model = cv2.face.LBPHFaceRecognizer_create()
multiple_model.read('multiple.xml')

# Modelo rostro
face_model = cv2.face.LBPHFaceRecognizer_create()
face_model.read('Face.xml')# Lectura del modelo entrenado

#Modelo forma de caminar
weight_path = 'C:/Users/almac/Desktop/GaitReid/model_sa_2_cnn.h5'
gait_model = GaitModel(weight_path=weight_path, num_classes=3)

cap = cv2.VideoCapture("Gait/GaitReid/videos/000/luis.mp4")

# Asignar pesos iniciales a las predicciones de cada modelo
weight_face = 0.4
weight_gait = 0.4
weight_softbiometric = 0.2

def combined_predictions(silueta_data, textura_data, color_data, face_data, gait_data):
def combined_softbiometric(result_silueta, result_textura, result_color):

```

Figura 56

Identificación de una persona con la fusión de los tres modelos generados, cámara 1

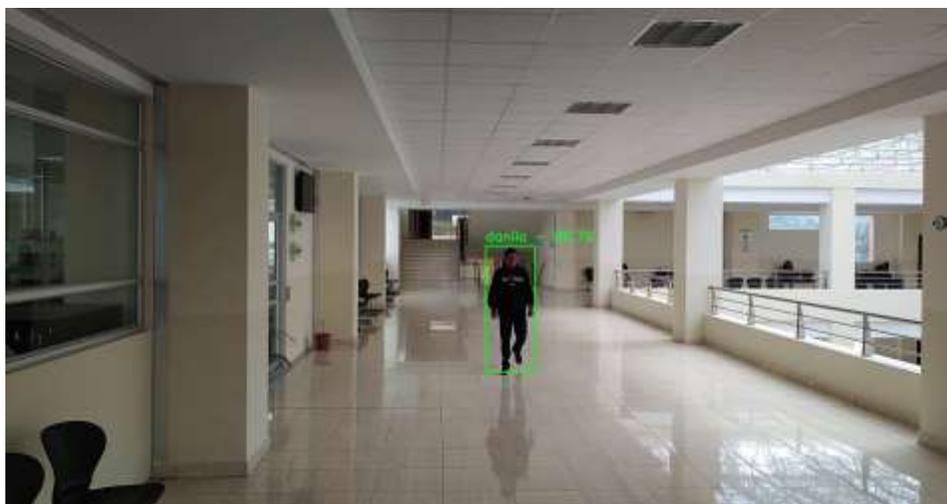


Figura 57

Identificación de una persona con la fusión de los tres modelos generados, cámara 2



Sprint 6: Validación, Ajustes e Implantación

El Sprint 6 resulta en la finalización del sistema de Re-Identificación de personas mediante técnicas de visión por computadora. Este Sprint se realizó entre el 13 de julio al 18 de agosto de 2023, previamente se han desarrollado todos los modelos que permiten el reconocimiento mediante características Biométricas (rostro y forma de caminar) y Soft-biométricas (silueta corporal, color y textura de la vestimenta), por ello esta fase consistió en garantizar que el sistema desarrollado Re-Identifique personas con la máxima precisión y eficacia posible.

Primeramente, se realizaron pruebas, en su respectivo entorno controlado, identificando cualquier fallo o área de mejora. Estas pruebas fueron importantes porque nos permitieron obtener datos sobre el rendimiento del sistema. Seguidamente, se procedió con la validación de los modelos fusionados, esto resulto en la obtención de datos que muestran el rendimiento del sistema con estándares de la industria, lo cual nos permitió determinar que el sistema de Re-identificación de personas se ajusta a los estándares de la industria.

Historia de Usuario 06: Validación, Ajustes e Implantación

La historia de usuario H.U.06 consiste en finalizar el proceso de implementación del sistema de Re-Identificación de personas, validando la eficacia y precisión de los modelos integrados, realizando ajustes necesarios y documentando todo el proceso y resultados obtenidos.

Detalle Historia de Usuario 06. La historia de usuario que está involucrada en este Sprint es la H.U.06, la cual se encuentra destallada en la Tabla 21.

Tabla 21

Historia de usuario para validación, ajustes e implantación

Historias de Usuario	
Numero: H.U.06	Usuario: Usuario
Nombre historia: Validación, Ajustes e Implantación	Numero Sprint: 6
Prioridad: Alta	Riesgo de desarrollo: Media
Puntos estimados (días): 25	interacción asignada: 1
Desarrolladores responsables: Bryan Chicaiza, Lenin Arteaga	
Descripción: Como usuario quiero que el sistema de Re-Identificación permita el reconocimiento de una persona a través de características biométricas (rostro y forma de caminar) y soft-biométricas (silueta corporal, color y textura de la vestimenta) mientras se procesa la primera entrada de video de la primera cámara de seguridad y se Re-Identifique en tiempo real a la persona mientras se procesa la segunda entrada de video procedente de la segunda cámara de seguridad.	
Validación:	

Historias de Usuario

- Se identifica a una persona mediante el modelo fusionado que combina características biométricas (rostro y forma de caminar) y soft-biométricas (silueta corporal, color y textura de la vestimenta) con la entrada de video correspondiente a la primera cámara de seguridad.
 - Se re-identifica a una persona con la entrada de video correspondiente a la segunda cámara de seguridad, teniendo un registro previamente obtenido por la primera cámara de seguridad.
 - Se evalúa la precisión del sistema aplicando métricas evaluación.
 - Se prueba y ajusta el sistema de Re-Identificación de personas para mejorar los resultados obtenidos.
-

Sprint Backlog 06

El Sprint 6, que se llevó a cabo del 13 de julio de 2023 al 18 de agosto de 2023, se centró en finalizar el proceso de implementación del sistema de Re-Identificación de personas, donde se evalúa la eficiencia y precisión de los modelos desarrollados previamente, así también, se realizan las pruebas pertinentes con los videos proporcionados por las cámaras de seguridad en tiempo real.

Las tareas que se realizaron durante el Sprint 6 se presentan en la Tabla 22.

Tabla 22

Sprint Backlog 06

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
H.U.06	Realizar pruebas para evaluar el funcionamiento del sistema en un	40	13/07/2023	17/07/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

HU ID	Tareas	Horas	Inicio	Fin	Responsable	Estado
	entorno controlado.					
H.U.06	Identificar posibles fallos o áreas de mejora, comparar los resultados.	56	18/07/2023	24/07/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.06	Realizar ajustes en los modelos e integración de estos.	56	25/07/2023	1/08/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado
H.U.06	Documentar todo el proceso de desarrollo en una documentación para su presentación	48	2/08/2023	18/08/2023	Bryan Chicaiza, Lenin Arteaga	Finalizado

Burndown chart sprint 06

El gráfico Burndown Chart del sprint 6 (ver Figura 58) compara el trabajo realizado con el planificado para este sprint. Este gráfico es importante para supervisar el progreso del equipo y cambiar el ritmo de trabajo si es necesario.

Las horas de trabajo pendientes se representan en el eje vertical, siendo el total previsto para el trabajo de este sprint de 200 horas. Los días de trabajo del sprint, del 13 de julio al 18 de agosto de 2023, se muestran en el eje horizontal. El gráfico muestra que durante la ejecución del sprint no presentaron mayores complicaciones.

Figura 58

Burndown chart sprint 06



Resultados Sprint 06

Al finalizar este sprint se obtuvo la arquitectura final del modelo CNN utilizada para la Re-Identificación de personas por la forma de caminar (ver Figura 59 y Figura 60).

Figura 59

Arquitectura del modelo CNN para la Re-Identificación de personas

```

class SpatialAttention(Layer):
    # ...

class GaitModel:
    def __init__(self, weight_path=None, num_classes=8):
        self.weight_path = weight_path
        self.model = None
        self.input_shape = None # default shape
        self.num_classes = num_classes

    def get_model(self):
        if self.model != None:
            return self.model

        # Define the model parameters
        # Number of classes (e.g., "normal walking" and "abnormal walking")
        # Create the model
        model = tf.keras.Sequential()
        # Add convolutional and pooling layers
        model.add(Conv2D(32, kernel_size=(3, 3), activation='relu', input_shape=self.input_shape, kernel_regularizer=l2(0.001)))
        model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))
        model.add(Conv2D(64, kernel_size=(3, 3), activation='relu', kernel_regularizer=l2(0.001)))
        model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))
        model.add(SpatialAttention())

```

Figura 60

Aumento de datos para mejorar la precisión del modelo por la forma de andar

```

datagen_train = ImageDataGenerator(
    rotation_range=20, # randomly rotate images in the range (degrees, 0 to 180)
    horizontal_flip=True) # randomly flip images
datagen_train.fit(X_train)
datagen_val = ImageDataGenerator()
datagen_val.fit(X_val)

epochs = 200
batch_size = 32

model_sa.compile(loss='categorical_crossentropy', optimizer=tf.keras.optimizers.Adam(), metrics=['accuracy'])

# entrenamos con el datagenerator (nos permite aplicar transformaciones en tiempo real)
model_sa.fit(datagen_train.flow(X_train, y_train, batch_size=batch_size),
            epochs=epochs,
            validation_data=datagen_val.flow(X_val, y_val, batch_size=batch_size))

```

Los ajustes en la arquitectura del modelo y el aumento de datos realizados a las imágenes de entrenamiento iniciales aplicando rotación vertical de hasta 20 y fotos reflejadas mejoraron la precisión del modelo. Estas mejoras se ven reflejadas en los resultados de la matriz de confusión obtenidas luego del entrenamiento del modelo (ver Tabla 23) y al aumento del porcentaje del 41,51% en comparación al primer modelo.

Tabla 23

Resultado de accuracy del primer y segundo modelo obtenidos al utilizar la cámara 2

Clasificador/	Clase	Métrica	Primer modelo	Segundo modelo
Descriptor			propuesto	propuesto
CNN/	Danilo	Accuracy	0.5000	1.000
GEI	Luis	Accuracy	0.5000	0.9166
	Pedro	Accuracy	0.5833	0.9166
PROMEDIO			0.5277 ± 0.0392	0.9440 ± 0.0448

Nota. Los resultados y modelos o propuestos se cubren con más detalle en el siguiente capítulo.

Capítulo IV

Validación del Sistema

En este capítulo se describen detalladamente las pruebas realizadas a los distintos modelos generados para la Re-Identificación de personas. El marco de evaluación, los parámetros utilizados y los resultados de las pruebas se presentan y analizan en las siguientes secciones de este capítulo.

Marco de Evaluación

La validación del sistema de Re-Identificación, fue realizada utilizando dos cámaras no superpuestas en un entorno controlado: la primera cámara reconoció rostros y adquirió el conjunto de datos para extraer las características de la forma de andar, color, textura y silueta utilizados para entrenar los modelos soft-biométricos; la segunda cámara sirvió para capturar imágenes donde el individuo se encuentre de espalda, empleados para el reconocimiento por las tres características soft-biométricas y la forma de andar.

El sistema utiliza la librería de Python multiprocessing que permite trabajar con tres procesadores simultáneamente. El primer procesador se encarga de buscar personas y reconocer rostros en las imágenes captadas por la primera cámara. A continuación, el segundo procesador extrae las características correspondientes al rostro, color, textura y silueta del conjunto de frames capturados por la cámara 1.

Las imágenes del promedio de las siluetas (Gait Energy Image - GEI) utilizados como entrada para el modelo de la forma de caminar también son extraídas por el segundo procesador. Los datos obtenidos se utilizan para extraer el conjunto de características para entrenar el modelo de la forma de caminar y el modelo soft-biométrico combinado respectivamente. El tercer procesador se encargó de realizar el proceso de Re-Identificación empleando los modelos previamente entrenados. Para realizar estas pruebas se utilizó como escenario el Corredor de la ESPE sede Latacunga.

Aplicación de métricas de evaluación

Las métricas utilizadas para la evaluación se enumeran en la sección "Métricas de evaluación" del capítulo 3. En particular, se utilizó la métrica de precisión porque evalúa todos los datos de la matriz de confusión (Hossin & Sulaiman, 2015).

Se utilizaron tres vídeos correspondientes a las clases Pedro, Luis y Danilo, para recopilar los datos necesarios para entrenar y probar el algoritmo. Los vídeos se grabaron en un entorno con iluminación adecuada y libre de oclusión. Para cada persona, se recogió un vídeo caminando de frente y sosteniendo una espada, filmado con la primera y la segunda cámara.

Las pruebas fueron realizadas utilizando tres procesadores para realizar simultáneamente las tareas de reconocimiento de personas y rostros, generación de modelos y Re-Identificación de personas. Al repartir las tareas entre los procesadores y utilizar sus recursos, se presentan mejoras en el rendimiento del sistema de implementado.

Pruebas del algoritmo de identificación facial, Re-Identificación forma de andar y Re-Identificación de color, textura y silueta.

Pruebas del algoritmo de identificación facial

En las pruebas se realizaron utilizando los modelos correspondientes a las tres categorías captadas por la cámara 1 (Luis, Danilo y Pedro). En este punto el sistema ya cuenta con mejoras implementadas para aumentar la precisión en las predicciones del modelo.

En este algoritmo se generaron modelos para la identificación de personas utilizando características biométricas del rostro, para las clases correspondientes de individuos grabados en vídeo, obtenidos de dos escenarios siguiendo una misma ruta. En la Tabla 24, Tabla 25 y Tabla 26 se muestran la matriz de confusión, los valores de precisión y los porcentajes de predicción respectivamente.

Tabla 24

Matriz de confusión del modelo de identificación facial de la clase Danilo

Cámara 1		
	POSITIVOS	NEGATIVOS
POSITIVOS	81 (VP)	0 (FP)
NEGATIVOS	0 (FN)	80 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Danilo, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 25

Matriz de confusión del modelo de identificación facial de la clase Luis

Cámara 1		
	POSITIVOS	NEGATIVOS
POSITIVOS	73 (VP)	5 (FP)
NEGATIVOS	6 (FN)	72 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Luis, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 26

Matriz de confusión del modelo de identificación facial de la clase Pedro

Cámara 1		
	POSITIVOS	NEGATIVOS
POSITIVOS	66 (VP)	14 (FP)

Cámara 1		
NEGATIVOS	3 (FN)	74 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Pedro, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

A pesar de haber empleado un entorno controlado, las condiciones de este escenario como el brillo y el reflejo del suelo influyeron en el número de personas detectadas por el modelo en los diferentes frames que compone la grabación. El número fijo establecido para el dataset fue de 300 imágenes por video.

Incluso en un entorno controlado, el brillo y la reflectividad del suelo afectan al número de personas que el modelo puede reconocer en los frames del video. Para este proyecto se establecieron un total de 300 imágenes como número de fotogramas almacenados por video.

Luego de haber obtenido los datos de la matriz de confusión se realiza el cálculo de porcentaje de exactitud de cada uno del modelo, aplicando la fórmula:

$$accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

La Tabla 27 muestra los resultados obtenidos tras aplicar la ecuación (1) para el cálculo de accuracy de las tres clases: Danilo, Luis y Pedro.

Tabla 27

Resultados de la exactitud del modelo de las 3 clases mediante la identificación facial

Clasificador/ Descriptor	Clase	Métricas	Valor
	Danilo	Accuracy	1.000

Clasificador/	Clase	Métricas	Valor
Descriptor			
Haar	Luis	Accuracy	0.9294
Cascade/	Pedro	Accuracy	0.8917
LBPH			

Una vez obtenida la métrica de precisión (Accuracy), se calcula el porcentaje de predicción. La Tabla 28 muestra la información para las tres clases, Daniel, Pedro y Luis, incluyendo clase, escenario, característica reconocida, predicción, resultados esperados y alcanzados, porcentaje de predicción y tasa media de reconocimiento tras la prueba.

Tabla 28

Porcentaje de predicción del modelo de identificación facial

Clase	Escenario	Característica biométrica	Predicción	Resultado esperado	Resultado Obtenido	Porcentaje de predicción
Danilo	Cámara 1	Rostro	Identificado	Danilo	Danilo	1.000
Luis	Cámara 1	Rostro	Identificado	Luis	Luis	0.9294
Pedro	Cámara 1	Rostro	Identificado	Pedro	Pedro	0.8917
PROMEDIO						0.9403 ± 0.0448

La tasa media de reconocimiento de caras fue del 94,03%, es decir, son reconocidas 94 de cada 100 caras reconocidas. El valor mínimo predicción del modelo está entre 89% y el 100% correspondientes al valor mínimo y máximo respectivamente.

Pruebas del algoritmo de Re-Identificación de color, textura y silueta

Par este algoritmo se generaron para la Re-Identificación de personas modelos utilizando características soft-biométricas: color, la textura y la silueta, para las clases correspondientes de

individuos grabados en vídeo, obtenidos de dos escenarios siguiendo una misma ruta. En la Tabla 29, Tabla 30 y Tabla 31 se muestran la matriz de confusión, los valores de precisión y los porcentajes de predicción respectivamente.

Tabla 29

Matriz de confusión del modelo combinado de la clase Danilo

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	225 (VP)	0 (FP)	189 (VP)	44 (FP)
NEGATIVOS	0 (FN)	227 (VN)	13 (FN)	190 (VN)

Nota. La tabla muestra los resultados de los verdaderos positivos (VP), falsos negativos (FN), verdaderos negativos (VN) y falsos positivos (FP) para la clase Danilo, se indica el resultado de la matriz de confusión.

Tabla 30

Matriz de confusión del modelo combinado de la clase Luis

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	190 (VP)	20 (FP)	193 (VP)	57 (FP)
NEGATIVOS	22 (FN)	71 (VN)	10 (FN)	190 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Luis, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 31

Matriz de confusión del modelo combinado de la clase Pedro

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	82 (VP)	6 (FP)	205 (VP)	34 (FP)
NEGATIVOS	19 (FN)	193 (VN)	21 (FN)	189 (VN)

Nota. La tabla muestra los resultados de los verdaderos positivos (VP), falsos negativos (FN), verdaderos negativos (VN) y falsos positivos (FP) para la clase Pedro, se indica el resultado de la matriz de confusión.

La Tabla 32 muestra los resultados obtenidos tras aplicar la ecuación (1) para el cálculo de accuracy de las tres clases: Danilo, Luis y Pedro.

Tabla 32

Resultados del modelo Soft-biométrico combinado

Clasificador/	Clase	Métricas	Valor
Descriptor			
SVM/	Danilo	Accuracy	0.8692
LBPH	Luis	Accuracy	0.8511
	Pedro	Accuracy	0.8775

Finalmente, luego de obtener los resultados de la métrica Accuracy se procede a realizar el cálculo de porcentaje de predicción. Los resultados obtenidos se muestran en la Tabla 33.

Tabla 33

Porcentaje de predicción del modelo soft-biométrico combinado

Clase	Escenario	Característica	Predicción	Resultado esperado	Resultado Obtenido	Porcentaje de predicción
		soft-biométrica				
Danilo	Cámara 2	Color, textura y silueta	Identificado	Danilo	Danilo	0.8692
Luis	Cámara 2	Color, textura y silueta	Identificado	Luis	Luis	0.8511
Pedro	Cámara 2	Color, textura y silueta	Identificado	Pedro	Pedro	0.8775
PROMEDIO						0.8659 ± 0.0110

El promedio obtenido para la identificación de modelo soft-biométrico combinado es del 86.59 %, lo que significa que de cada 100 individuos puede reconocerá 86, por lo que el rango de predicción de este modelo esta entre el 85.11 % como valor mínimo y con valor máximo de predicción del 87.75 %.

Pruebas del algoritmo de Re-Identificación forma de andar

A continuación, se presenta la matriz de confusión, valor del Accuracy y el porcentaje de predicción del modelo generado para la Re-Identificación a través de la característica biométrica forma de andar, de cada una de las clases correspondientes a los individuos capturados en video, los cuales fueron obtenidos en dos escenarios siguiendo una ruta (ver Tabla 34, Tabla 35 y Tabla 36).

Tabla 34

Matriz de confusión del modelo de la forma de andar de la clase Danilo

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	6 (VP)	0 (FP)	6 (VP)	0 (FP)

	Cámara 1		Cámara 2	
NEGATIVOS	0 (FN)	6 (VN)	3 (FN)	3 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Danilo, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 35

Matriz de confusión del modelo de la forma de andar de la clase Luis

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	6 (VP)	0 (FP)	3 (VP)	3 (FP)
NEGATIVOS	0 (FN)	6 (VN)	0 (FN)	6 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Luis, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 36

Matriz de confusión del modelo de la forma de andar de la clase Pedro

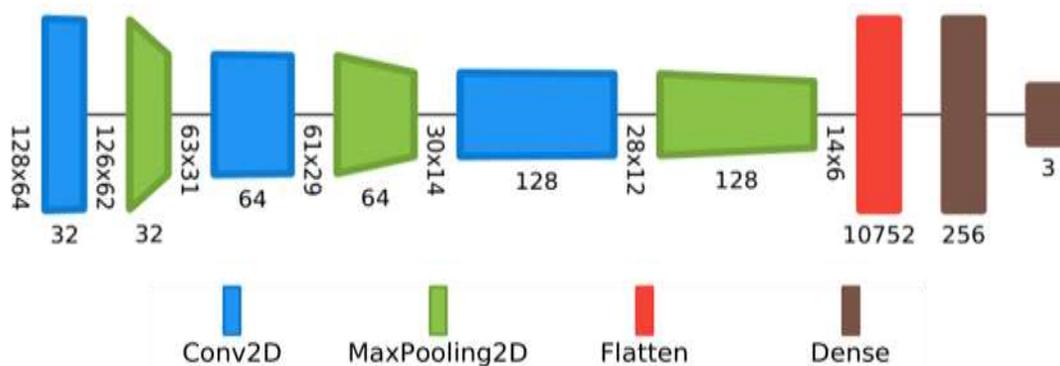
	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	6 (VP)	0 (FP)	4 (VP)	2 (FP)
NEGATIVOS	0 (FN)	6 (VN)	3 (FN)	3 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Pedro, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

La Tabla Tabla 37 muestra los resultados obtenidos tras aplicar la ecuación (1) para el cálculo de accuracy de las tres clases: Danilo, Luis y Pedro. La exactitud obtenida al utilizar como datos de entrada las imágenes de la primera cámara donde el individuo se visualiza de frente presenta un 100 %, sin embargo, al introducir los datos al modelo cuando la persona está de espalda, capturados por la segunda cámara, este primer modelo (ver Figura 61) presenta errores sin tener la capacidad de predecir correctamente, manteniendo una tendencia hacia la clase Danilo (ver Figura 62).

Figura 61

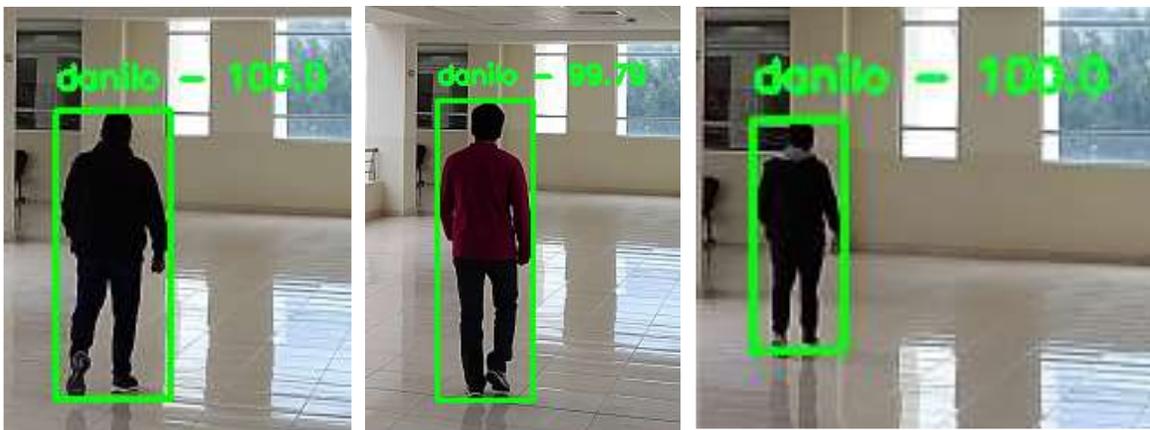
Primera arquitectura CNN propuesta para la Re-Identificación de personas por la forma de caminar



Nota. La figura muestra la primera arquitectura empleada para el proceso de Re-Identificación por la forma de andar aplicando lo definido en el Marco de evaluación. Se utilizó el optimizador Adam con una tasa de aprendizaje de 0,0001. La red se entrenó con un batch size de 30 con 200 epochs.

Figura 62

Resultados del primer modelo utilizado para la Re-Identificación de personas por la forma de caminar



Resultado esperado:

Danilo

a)

Resultado esperado:

Luis

b)

Resultado esperado:

Pedro

c)

Nota. La figura muestra los errores en la predicción del modelo de la forma de andar sin variaciones en su estructura, entradas e hiperparámetros. En las secciones a), b) y c) de la imagen se presenta que todas las entradas son predichas como Danilo. La sección a) es correcta, sin embargo, tanto b) y c) son incorrectos.

Tabla 37

Resultados de la exactitud del modelo combinado

Clasificador/ Descriptor	Clase	Métricas	Valor
CNN/	Danilo	Accuracy	0.5000
GEI	Luis	Accuracy	0.5000
	Pedro	Accuracy	0.5833

Nota. Resultado de exactitud del modelo de la forma de andar ingresando como entrada los datos (GEI) de las imágenes capturadas por la cámara 2 de la clase Danilo, Luis y Pedro.

Luego de obtener los resultados de la métrica Accuracy se procede a realizar el cálculo de porcentaje de predicción. En la Tabla 43 se muestra la clase, escenario, característica a identificar, predicción, resultado esperado y obtenido, su porcentaje de predicción y el promedio de identificación obtenidas tras la ejecución de las pruebas de las tres clases: Danilo, Pedro y Luis.

Tabla 38

Porcentaje de predicción del modelo de la forma de andar

Clase	Escenario	Característica biométrica	Predicción	Resultado esperado	Resultado Obtenido	Porcentaje de predicción
Danilo	Cámara 2	Forma de andar	Identificado	Danilo	Danilo	0.5000
Luis	Cámara 2	Forma de andar	No identificado	Luis	Danilo	0.5000
Pedro	Cámara 2	Forma de andar	No identificado	Pedro	Danilo	0.5880
PROMEDIO						0.5293 ± 0.0414

El promedio obtenido para la identificación de la forma de caminar es del 52.93 %, quedando fuera del rango aceptable en la literatura. Para solventar este problema se agregó realizó modificaciones a la arquitectura del modelo. A esta arquitectura se agregó una capa de atención espacial para que el modelo aprenda a enfocarse en regiones específicas de la imagen, y asigne diferentes niveles de importancia a las distintas regiones de la imagen de entrada. Además, se agregó una capa de dilución (Dropout) y regularización L2 para reducir el sobreajuste del modelo (ver Figura 63).

Dado que se disponía únicamente de imágenes de la parte frontal de la persona para entrenar el modelo, se realizó un aumento de datos para otorgar varianza al conjunto disponible. Al dataset de entrenamiento se le aplicó giros verticales de hasta 20 grados y se invirtieron las imágenes. Luego de

entrenar el modelo, se obtuvieron menos falsos positivos y falsos negativos en la matriz de confusión de las tres clases (ver Tabla 39, Tabla 40 y Tabla 41), en consecuencia, mejoró los resultados de la métrica accuracy en un 41.51%, en comparación al primer modelo propuesto (ver Tabla 43).

Tabla 39

Matriz de confusión del nuevo modelo de la forma de caminar de la clase Danilo

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	6 (VP)	0 (FP)	6 (VP)	0 (FP)
NEGATIVOS	0 (FN)	6 (VN)	0 (FN)	6 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Danilo, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 40

Matriz de confusión del nuevo modelo de la forma de caminar combinado de la clase Luis

	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	6 (VP)	0 (FP)	6 (VP)	0 (FP)
NEGATIVOS	0 (FN)	6 (VN)	1 (FN)	5 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Luis, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

Tabla 41

Matriz de confusión del nuevo modelo de la forma de caminar de la clase Pedro

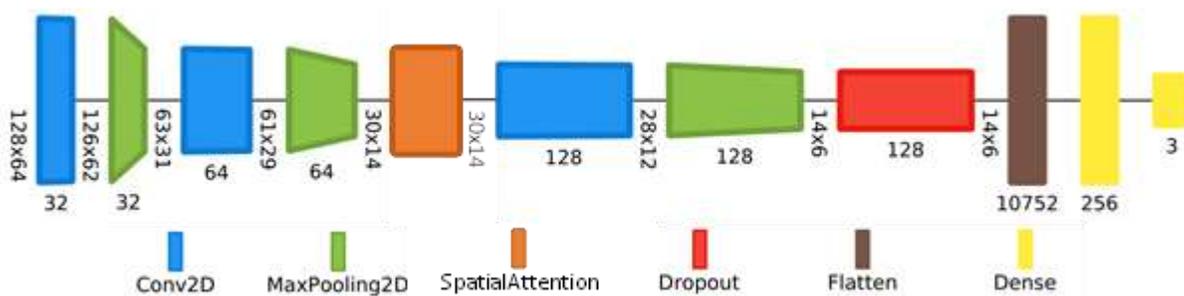
	Cámara 1		Cámara 2	
	POSITIVOS	NEGATIVOS	POSITIVOS	NEGATIVOS
POSITIVOS	6 (VP)	0 (FP)	5 (VP)	1 (FP)
NEGATIVOS	0 (FN)	6 (VN)	0 (FN)	6 (VN)

Nota. La tabla muestra los resultados de la matriz de confusión obtenidos de la clase Pedro, donde VP representa los Verdaderos Positivos, FN Falsos Negativos, VN Verdaderos Negativos y FP Falsos Positivos.

La Tabla 42 muestra los resultados obtenidos tras aplicar la ecuación (1) para el cálculo de accuracy de las tres clases: Danilo, Luis y Pedro. Utilizando el nuevo modelo (ver Figura 62) con las modificaciones anteriormente explicadas, este tiene la capacidad de diferenciar entre las tres diferentes clases (ver Figura 64).

Figura 63

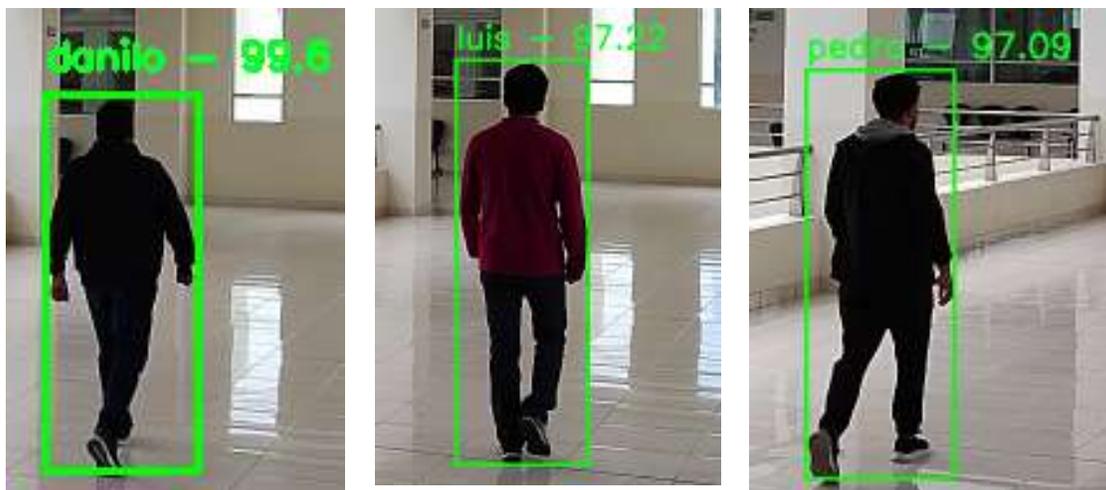
Arquitectura CNN modificada, propuesta para la Re-Identificación de personas por la forma de caminar



Nota. La figura muestra la arquitectura modificada, empleada para el proceso de Re-Identificación por la forma de caminar aplicando lo definido en el Marco de evaluación. Se utilizó el optimizador Adam con una tasa de aprendizaje de 0,0001, además, a cada capa Convolución se agregó la regularización de peso L2 de valor 0.01 y un valor de 0.5 para la capa Dropout. La red se entrenó con un batch size de 32 con 200 epochs.

Figura 64

Resultados del nuevo modelo utilizado para la Re-Identificación de personas por la forma de caminar



Resultado esperado:

Danilo

a)

Resultado esperado:

Luis

b)

Resultado esperado:

Pedro

c)

Nota. La figura muestra los resultados de predicción del modelo modificado. En las secciones a), b) y c) de la imagen los resultados predichos por el modelo son los esperados.

Tabla 42

Resultados de la exactitud del modelo combinado

Clasificador /	Clase	Métricas	Valor
Descriptor			
CNN/GEI	Danilo	Accuracy	1.0000
	Luis	Accuracy	0.9166
	Pedro	Accuracy	0.9166

Nota. Resultado de exactitud del modelo combinado ingresando como entrada los datos (GEI) de la clase Danilo, Luis y Pedro.

A continuación, en la Tabla 43 se muestra la clase, escenario, característica a identificar, predicción, resultado esperado y obtenido, su porcentaje de predicción y el promedio de identificación obtenidas tras la ejecución de las pruebas de las tres clases: Danilo, Pedro y Luis.

Tabla 43

Porcentaje de predicción del modelo de la forma de caminar

Clase	Escenario	Característica biométrica	Predicción	Resultado esperado	Resultado Obtenido	Porcentaje de predicción
Danilo	Cámara 2	Forma de caminar	Identificado	Danilo	Danilo	1.0000
Luis	Cámara 2	Forma de caminar	Identificado	Luis	Luis	0.9166
Pedro	Cámara 2	Forma de caminar	Identificado	Pedro	Pedro	0.9166
PROMEDIO						0.9444 ± 0.0393

El promedio obtenido para la identificación de la forma de caminar es del 94.44 %, lo que significa que de cada 100 individuos puede reconocerá 94. El resultado obtenido se encuentra en el rango aceptable de la literatura. Sin embargo, debido a las limitaciones presentadas por el número de cámaras disponibles para extraer los datos de entrenamiento, los sistemas de identificación de personas por la forma de caminar que utilizan un único ángulo de cámara pueden ser más susceptibles de errores en comparación a los sistemas que utilizan un único ángulo de cámara (Hossin & Sulaiman, 2015; Wang *et al.*, 2011). Esto es debido a que estos sistemas solo pueden aprender una parte de la forma de la caminar, lo que puede dificultar la identificación de la persona.

Capítulo V

Conclusiones

Finalmente, los resultados de la investigación se resumen de la siguiente manera:

- La investigación realizada en diversas bases de datos, incluyendo IEEE, WOS, Scopus y Google Scholar, permitió conocer diversas técnicas avanzadas de Visión por Computador, modelos de Machine Learning y/o algoritmos utilizados para la creación de sistemas de Re-Identificación de personas utilizando características soft-biométricas (textura, color y silueta) y biométricas (rostro y marcha). Estas técnicas fueron utilizadas para generar los diferentes modelos utilizados para el funcionamiento del sistema implementado.
- El desarrollo del proyecto fue realizado empleando el marco de trabajo SCRUM, puesto que esta metodología está diseñada para ser flexible y permitir ajustes en los requisitos. Esto permitió generar historias de usuario y la estimación de tiempos de las diferentes tareas considerando que puede haber cambio, debido a la complejidad que presenta los sistemas de Re-Identificación.
- Al final se obtuvo un total de 600 imágenes de la silueta de las personas por cada video, del cual se puede generar un total de 6 imágenes que representan el promedio de las siluetas utilizadas para el entrenamiento del modelo la forma de caminar y 300 imágenes tanto del rostro como de la persona para entrenar el modelo para la detección facial y el soft-biométrico combinado respectivamente.
- Los principales retos que dificultaron obtener datos de buena calidad fue el brillo de los objetos, reflejo de las baldosas y ventanales en el entorno por lo cual para solventar estos problemas se realizaron ajustes a los diferentes modelos y mejorar los datos para el entrenamiento. Luego de aplicar estas modificaciones se obtuvo como resultado del sistema de Re-Identificación de personas utilizando características biométricas (el rostro y la forma de caminar) y soft-

biométricas (silueta corporal, textura y color de la ropa) una precisión del 91 %, ubicándose entre el 70% al 99% por lo que se encuentra dentro del rango mínimo y máximo encontrado en la literatura científica (Nithyakani et al., 2019).

- Utilizar Haar cascade en la identificación de personas por el rostro ha demostrado ser útil en la extracción de características faciales discriminativas al conjunto de datos de entrada proporcionado. Por otra parte, SVM resultó ser útil para separar las diferentes clases y encontrar patrones en la gran cantidad atributos que presentan los datos empleados para la Re-Identificación por características soft-biométricas (color, textura y silueta). CNN a su vez ha mostrado ser útil para aprender características invariantes al ángulo de captura (en cierta medida) y generalizar mejor a diferentes condiciones de captura disminuyendo la complejidad de los datos necesarios para entrenar el modelo.
- Al combinar las características soft-biométricas se obtuvieron mejores resultados en la predicción en comparación a utilizarlos de forma independiente, puesto que se pudo obtener un conjunto de características más representativas de un individuo. Esta mejora fue de gran aporte para el sistema puesto que fue principalmente utilizado para Re-Identificar a la persona cuando no puede detectar el rostro y en las imágenes del video no se puede captar el movimiento de los pies durante la marcha donde ninguna de las características biométricas podía aportar a porcentaje de predicción.
- En la literatura comúnmente se hace uso del promedio de las siluetas, sin embargo, por lo general los modelos son entrenados con un conjunto de datos obtenidos en diferentes ángulos de la cámara. Por lo tanto, los modelos utilizados por los diferentes estudios sin modificaciones no solventaban los requerimientos de este estudio. Una de las soluciones encontrada para resolver este inconveniente fue agregar una capa de atención espacial para que aprenda a

enfocarse en regiones específicas de la imagen y se pueda obtener características más relevantes de la forma de caminar.

- Los sistemas de Re-Identificación de la forma de andar que utilizan un solo ángulo de cámara pueden dar lugar a errores, ya que el modelo pierde información clave de la marcha que son captados desde otro ángulo de vista del individuo. Ante esta problemática se integran otras características como son las soft-biométricas que complementan a estos sistemas, permitiendo que estos puedan ser aplicados en videovigilancia, seguridad pública y prevención del delito.
- Al utilizar un único ángulo de cámara para generar el conjunto de datos utilizado para el entrenamiento, los datos de entrenamiento del modelo se limitan al ángulo de la cámara y se pierde información importante sobre la marcha en relación con la perspectiva del usuario respecto a la cámara. Por tanto, en futuros trabajos, proponemos el uso de cámaras de un circuito cerrado para captar más información sobre la forma de caminar y crear un modelo más fiable ante estos cambios.

Bibliografía

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. *USENIX*, 21.
- Abderrahmane, E., & Haj Thami, R. O. (2013). *Conception and development of a video surveillance system for detecting, tracking and profile analysis of a person | IEEE Conference Publication | IEEE Xplore*. <https://ieeexplore.ieee.org/document/6728128>
- Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12), 2037–2041. <https://doi.org/10.1109/TPAMI.2006.244>
- Bakchy, S. C., Islam, M. R., Mahmud, M. R., & Imran, F. (2022). *Human Gait Analysis using Gait Energy Image* (arXiv:2203.09549). arXiv. <https://doi.org/10.48550/arXiv.2203.09549>
- Chen, Y., Zhu, X., & Gong, S. (2017). Person Re-identification by Deep Learning Multi-scale Representations. *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, 2590–2600. <https://doi.org/10.1109/ICCVW.2017.304>
- Chugchilan Hinojosa, W. A., & Guanochanga Farfán, J. H. (2023). *Desarrollo de un Sistema de re-identificación de personas a través de las características faciales y softbiométricas el color de la vestimenta, y modelos holísticos y no holísticos, utilizando algoritmos de Visión por Computadora y Machine Learning* [BachelorThesis, Carrera de Ingeniería en Software. Universidad de las Fuerzas Armadas ESPE. Extensión Latacunga.]. <http://repositorio.espe.edu.ec/jspui/handle/21000/35732>
- Cohn, M. (2009). *Succeeding with Agile: Software Development Using Scrum* (1st edition). Addison-Wesley Professional.

- Cortes, C., & Vapnik, V. (1995a). Support-vector networks. *Machine Learning*, 20(3), 273–297.
<https://doi.org/10.1007/BF00994018>
- Cortes, C., & Vapnik, V. (1995b). Support-vector networks. *Machine Learning*, 20(3), 273–297.
<https://doi.org/10.1007/BF00994018>
- Cuzzolin, F. (2010). *Multilinear Modeling for Robust Identity Recognition from Gait* [Chapter]. Behavioral Biometrics for Human Identification: Intelligent Applications; IGI Global.
<https://doi.org/10.4018/978-1-60566-725-6.ch008>
- Dalal, N., & Triggs, B. (2005a). Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 1, 886–893 vol. 1. <https://doi.org/10.1109/CVPR.2005.177>
- Dalal, N., & Triggs, B. (2005b). Histograms of Oriented Gradients for Human Detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 1, 886–893. <https://doi.org/10.1109/CVPR.2005.177>
- Dantcheva, A., Velardo, C., D'Angelo, A., & Dugelay, J.-L. (2011). Bag of soft biometrics for person identification: New trends and challenges. *Multimedia Tools and Applications*, 51.
<https://doi.org/10.1007/s11042-010-0635-7>
- Das, P., Das, D., & Saharia, S. (2014). Gait Analysis and Recognition for Human Identification. *International Journal of Electronics and Applied Research (IJEAR)*, 2014, 45–54.
- Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9), 1627–1645. <https://doi.org/10.1109/TPAMI.2009.167>
- Gafurov, D. (2011). Person re-identification using biometrics: Challenges, approaches, and applications. *In Biometrics*, 175–198.

- Goffredo, M. (2014). Gait analysis for human identification based on silhouette-based GEI. *International Journal of Image and Data Fusion*, 5(3), 231–250.
- Gong, S., Cristani, M., Loy, C. C., & Hospedales, T. M. (2014). The Re-identification Challenge. In S. Gong, M. Cristani, S. Yan, & C. C. Loy (Eds.), *Person Re-Identification* (pp. 1–20). Springer London.
https://doi.org/10.1007/978-1-4471-6296-4_1
- Han, J., & Bhanu, B. (2006). Individual Recognition Using Gait Energy Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28, 316–322. <https://doi.org/10.1109/TPAMI.2006.38>
- Hossin, M., & Sulaiman, M. N. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 01–11.
<https://doi.org/10.5121/ijdkp.2015.5201>
- How Face Recognition Technology and Person Re-identification Technology Can Help Make Our World Safer and Secure | NEC.* (n.d.). Retrieved August 21, 2023, from
<https://www.nec.com/en/global/techrep/journal/g18/n02/180215.html>
- Hua, C., Zhao, X., Meng, W., & Pan, Y. (2022). Deep Person Re-identification with the Combination of Physical Biometric Information and Appearance Features. In Z. Qian, M. A. Jabbar, & X. Li (Eds.), *Proceeding of 2021 International Conference on Wireless Communications, Networking and Applications* (pp. 874–887). Springer Nature. https://doi.org/10.1007/978-981-19-2456-9_89
- Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20.
<https://doi.org/10.1109/TCSVT.2003.818349>
- Lawrence, S., Giles, C. L., Ah Chung Tsoi, & Back, A. D. (1997). Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks*, 8(1), 98–113.
<https://doi.org/10.1109/72.554195>

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), Article 7553.

<https://doi.org/10.1038/nature14539>

Lignia Tuapanta, L. D., & Pichucho Tandalla, R. V. (2023, February 22). *Sistema de Re-Identificación de personas a través de las características faciales y soft-biométricas la textura de la vestimenta, y modelos holísticos y no holísticos, utilizando algoritmos de Visión por Computadora y Machine Learning*. <https://repositorio.espe.edu.ec/handle/21000/35825?locale=es>

Mandar, S., Srinivas, S., Ramaswamy, S., Raman, V., & Hemnani, P. (2021). Human Re-Identification Using Gait Analysis. *2021 International Conference on Intelligent Technologies (CONIT)*, 1–4.

<https://doi.org/10.1109/CONIT51480.2021.9498427>

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Merusi, D. (1995). *Software Implementation Techniques, Second Edition: Writing Software in OpenVMS, OS/2, UNIX and Windows NT* (2nd edition). Digital Press.

Ming, Z., Zhu, M., Wang, X., Zhu, J., Cheng, J., Gao, C., Yang, Y., & Wei, X. (2022a). *Deep learning-based person re-identification methods: A survey and outlook of recent works* (arXiv:2110.04764).

arXiv. <http://arxiv.org/abs/2110.04764>

Ming, Z., Zhu, M., Wang, X., Zhu, J., Cheng, J., Gao, C., Yang, Y., & Wei, X. (2022b). *Deep learning-based person re-identification methods: A survey and outlook of recent works* (arXiv:2110.04764).

arXiv. <http://arxiv.org/abs/2110.04764>

Moctezuma-Ochoa, D. A. (2016). Re-identificación de personas a través de sus características soft-biométricas en un entorno multi-cámara de video-vigilancia. *Ingeniería. Investigación y Tecnología, VXII(2)*, 257–271.

Mogan, J. N., Lee, C. P., Lim, K. M., & Muthu, K. S. (2022). Gait-ViT: Gait Recognition with Vision

Transformer. *Sensors (Basel, Switzerland)*, 22(19), 7362. <https://doi.org/10.3390/s22197362>

- Mostafaie, F., Nabizadeh, Z., Karimi, N., & Samavi, S. (2020). A General Framework for Saliency Detection Methods. *2020 International Conference on Machine Vision and Image Processing (MVIP)*, 1–5. <https://doi.org/10.1109/MVIP49855.2020.9116881>
- Nambiar, A., Bernardino, A., & Nascimento, J. C. (2019). Gait-based Person Re-identification: A Survey. *ACM Computing Surveys*, *52*(2), 33:1-33:34. <https://doi.org/10.1145/3243043>
- Nasaruddin, N., Muchtar, K., Afdhal, A., & Dwiyanoro, A. P. J. (2020). Deep anomaly detection through visual attention in surveillance videos. *Journal of Big Data*, *7*(1), 87. <https://doi.org/10.1186/s40537-020-00365-y>
- Navajo, P. (2023, May 21). Un recorrido por Scrum. *Iniciativa Social*. <https://www.iniciativasocial.net/?p=1496>
- Nithyakani, P., Shanthini, A., & Ponsam, G. (2019). Human Gait Recognition using Deep Convolutional Neural Network. *2019 3rd International Conference on Computing and Communications Technologies (ICCCT)*, 208–211. <https://doi.org/10.1109/ICCCT2.2019.8824836>
- Nixon, M., Carter, J., Cunado, D., Huang, P., & Stevenage, S. (2006). Automatic Gait Recognition. In *Motion Analysis and Tracking* (Vol. 103, pp. 231–249). https://doi.org/10.1007/0-306-47044-6_11
- Niyogi & Adelson. (1994). Analyzing and recognizing walking figures in XYT. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*, 469–474. <https://doi.org/10.1109/CVPR.1994.323868>
- Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, *29*(1), 51–59. [https://doi.org/10.1016/0031-3203\(95\)00067-4](https://doi.org/10.1016/0031-3203(95)00067-4)

- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002a). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002b). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
- Osuna, E., Freund, R., & Girosit, F. (1997). Training support vector machines: An application to face detection. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 130–136. <https://doi.org/10.1109/CVPR.1997.609310>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*, 6.
- Perea-Ortega, J., Montejo-Ráez, A., Martín-Valdivia, M., & López, L. (2009). Using Support Vector Machines as Learning Algorithm for Video Categorization. 6242, 373–376. https://doi.org/10.1007/978-3-642-15751-6_48
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016a). *You Only Look Once: Unified, Real-Time Object Detection* (arXiv:1506.02640). arXiv. <https://doi.org/10.48550/arXiv.1506.02640>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016b). *You Only Look Once: Unified, Real-Time Object Detection* (arXiv:1506.02640). arXiv. <https://doi.org/10.48550/arXiv.1506.02640>
- Rola, P., & Kuchta, D. (2015). Implementing Scrum Method in International Teams—A Case Study. *Open Journal of Social Sciences*, 3(7), Article 7. <https://doi.org/10.4236/jss.2015.37043>
- Shahbaz, A., Hoang, V.-T., & Jo, K.-H. (2019). Convolutional Neural Network based Foreground Segmentation for Video Surveillance Systems. *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, 1, 86–89. <https://doi.org/10.1109/IECON.2019.8927776>

- Sharma, P. (2019). A survey on color space and color model. *International Journal of Computer Science and Mobile Computing*, 8(1), 27–32.
- Shi, L.-F., Liu, Z.-Y., Zhou, K.-J., Shi, Y., & Jing, X. (2023). Novel Deep Learning Network for Gait Recognition Using Multimodal Inertial Sensors. *Sensors*, 23(2), Article 2.
<https://doi.org/10.3390/s23020849>
- Sun, Y., Wang, X., & Tang, X. (2014a). Deep Learning Face Representation from Predicting 10,000 Classes. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1891–1898.
<https://doi.org/10.1109/CVPR.2014.244>
- Sun, Y., Wang, X., & Tang, X. (2014b). *Deep Learning Face Representation from Predicting 10,000 Classes*. 1891–1898. <https://doi.org/10.1109/CVPR.2014.244>
- Turk, M., & Pentland, A. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86. <https://doi.org/10.1162/jocn.1991.3.1.71>
- Tyagi, M. (2021, July 24). *HOG(Histogram of Oriented Gradients)*. Medium.
<https://towardsdatascience.com/hog-histogram-of-oriented-gradients-67ecd887675f>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.
- Varona, J., González, J., Rius, I., & Villanueva, J. J. (2008). *Importance of detection for video surveillance applications*. <https://doi.org/10.1117/1.2965548>
- Vigneshwaran, T., Guruprakash, K. S., Nasreen, K. T., & Supraja, M. (2020). Effective Framework for real time video face recognition system. *Journal of Advanced Research in Dynamical and Control System*, 12(6), 680–684.
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 1, I–I. <https://doi.org/10.1109/CVPR.2001.990517>

- Viola, P., & Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2), 137–154. <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>
- Vishwakarma, D. K., & Upadhyay, S. (2018). A Deep Structure of Person Re-Identification Using Multi-Level Gaussian Models. *IEEE Transactions on Multi-Scale Computing Systems*, 4(4), 513–521. <https://doi.org/10.1109/TMSCS.2018.2870592>
- Wang, J., Wang, H., & Zhang, Y. (2011). Gait recognition using gait energy image and convolutional neural network. *Pattern Recognition*, 34(1), 186–195. <https://doi.org/10.1016/j.patrec.2010.11.033>
- Ye, M., Zhang, Q., Wang, L., Zhu, J., Yang, R., & Gall, J. (2013). A Survey on Human Motion Analysis from Depth Data (pp. 149–187). https://doi.org/10.1007/978-3-642-44964-2_8
- Zayat, W., & Senvar, O. (2020). Framework Study for Agile Software Development Via Scrum and Kanban. *International Journal of Innovation and Technology Management*, 17. <https://doi.org/10.1142/S0219877020300025>
- Zhang, B., Gao, Y., Zhao, S., & Liu, J. (2010). Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 19(2), 533–544. <https://doi.org/10.1109/TIP.2009.2035882>
- Zhao, R., Ouyang, W., & Wang, X. (2013). *Unsupervised Saliency Learning for Person Re-identification*. 3586–3593. <https://doi.org/10.1109/CVPR.2013.460>
- Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35(4), 399–458. <https://doi.org/10.1145/954339.954342>
- Zheng, L., Yang, Y., & Hauptmann, A. G. (2016). *Person Re-identification: Past, Present and Future* (arXiv:1610.02984). arXiv. <https://doi.org/10.48550/arXiv.1610.02984>

Anexos