UNIVERSIDAD NACIONAL SAN AGUSTIN DE AREQUIPA

FACULTAD DE INGENIERÍA DE PRODUCCIÓN Y SERVICIOS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

IDENTIFICACIÓN BIOMÉTRICA ÚNICA PARA HABILITAR LA INTEROPERABILIDAD EN EL PERÚ UTILIZANDO DEEP LEARNING Y BIGDATA

Tesis presentada por:

JAIME BENAVIDES ESQUIVEL

Para optar el Título Profesional de:

INGENIERO DE SISTEMAS

Asesor:

M. Sc. PERCY OSCAR HUERTAS NIQUEN

Arequipa – Perú

2022

Dedicatoria

*Esta **Tesis** está dedicado a mis padres, **Felipa Esquivel y Guillermo Benavides** por toda la paciencia, confianza y apoyo que tuvieron durante el proceso de desarrollo de esta **Tesis**.*

Agradecimientos

A mis Padres que con su apoyo me dieron la mejor herencia que un hijo puede tener, mi educación.

A mi hijo Emiliano y a mi esposa Yeisy que siempre me han estado compartiendo muchas alegrías y sabios consejos.

A mis Hermanos, Sandra, Aldo, Carlos, Dina y Roxana, que siempre me han estado compartiendo consejos y sus experiencias.

A todos mis profesores que, durante mi estancia en la UNSA, me apoyaron con sus conocimientos y experiencias, así aprendí y fortalecí mis conocimientos para enfrentar los retos de la sociedad, y de un mundo globalizado.

A mi asesor M. Sc. Percy Oscar Huertas Niquén, por su constante orientación académica, amistad y guía a lo largo del desarrollo del presente trabajo.

A mi Universidad y Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional de San Agustín por darme la formación académica necesaria para desarrollar esta tesis, a todos y al personal administrativo que siempre estuvo para apoyarme.

Presentación

SEÑOR DECANO DE LA FACULTAD DE INGENIERÍA DE PRODUCCIÓN Y SERVICIOS

SEÑOR DIRECTOR DE LA ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

De conformidad con las disposiciones vigentes del Reglamento de Grados y Títulos de la Facultad de Ingeniería de Producción y Servicios, y de la Escuela Profesional de Ingeniería de Sistemas pongo a vuestra disposición el presente TRABAJO DE INVESTIGACIÓN denominado "IDENTIFICACIÓN BIOMÉTRICA ÚNICA PARA HABILITAR LA INTEROPERABILIDAD EN EL PERÚ UTILIZANDO DEEP LEARNING Y BIGDATA" cuya aprobación me permitirá obtener el Título Profesional de Ingeniero de Sistemas.

Nombre y Apellidos
Bachiller de Ingeniería de Sistemas

Dictamen del Asesor

RESUMEN - PALABRAS CLAVES

El campo del reconocimiento facial ha experimentado avances constantes en las dos últimas décadas, hasta llegar a los precisos sistemas de reconocimiento en tiempo real disponibles hoy en día en los teléfonos móviles y en las redes sociales, Actualmente la inteligencia artificial ha evolucionado tanto, que ahora existe hasta como servicio Amazon Rekognition[1], así como Face ID[2] de Apple.

Esta tesis propone una solución de identidad biométrica única, una solución importante en nuestro país, puesto que nos va permitir identificar a una persona en entornos digitales, verificando digitalmente a la persona quien dice ser, y con esta identificación poder realizar trámites en línea, sin la necesidad de ir personalmente a una oficina de gobierno, habilitando así una sociedad y una economía digital y dejando una trazabilidad o huella digital.

La identificación de la identidad biométrica única, es una solución importante para que un ciudadano que vive en cualquier parte del país, pueda realizar un trámite en línea y de forma remota, utilizando video cámaras con lo que ayudara al funcionario de gobierno atenderlo en línea y de forma remota, ahorrando costos y tiempos en atención utilizando la ventanilla única digital.

El estado peruano ha planteado una ley para la identidad digital única, y tiene como objetivo permitir la autenticación en línea de la identidad de las personas peruanas y extranjeras cuando necesiten acceder a los servicios digitales brindados por las entidades públicas, esta ley fue publicado en septiembre del 2018, con el decreto legislativo nº 1412, en el capítulo II identidad digital, lo cual impulsa la implementación de la plataforma nacional de identificación y autenticación de la identidad digital.

La tesis presentada tiene un alcance propuesto desde el registro biométrico, y la generación de la credencial de la identidad digital, así como los mecanismos de autenticación y autorización digital, utilizando reconocimiento biométrico y firmas digitales, para habilitar una interoperabilidad digital [5] utilizando X-Road[3].

[1] Amazon Rekognition, La API de Rekognition y el servicio identificará objetos, personas, texto, escenas y actividades, además de detectar contenido inapropiado. https://aws.amazon.com/es/rekognition/.
[2] Face ID, autenticación mediante reconocimiento facial, Apple https://support.apple.com/en-us/HT208109.
[3] X-Road, Plataforma de Interoperabilidad, https://x-road.global/architecture

Esta tesis propone una solución de identidad biométrica única utilizando bibliotecas de identificación de rostro con OpenCV [1], así como para el reconocimiento de rostros utilizando técnicas de aprendizaje profundo, y aprendizaje por transferencia. Se realizó comparaciones con las diferentes arquitecturas VGG16[2], VGG19[2] y RESTNET50 [3] generando una precisión de entrenamiento mayor al 98%, trabajando con 9 clases y cada clase con 300 imágenes de 224x224 pixeles por imagen de entrada para cada clase, apoyados de las bibliotecas de TensorFlow, Keras, Sklearn, y validando el desempeño de los modelos utilizando la matriz de confusión, así como para su almacenamiento de la trazabilidad de identificación y autenticación en BigData, y la utilización de contenedores (kubernetes/docker) para tener un nivel de escalamiento nacional y global. Cada una de estas bibliotecas y la propuesta de solución son 100% software libre.

Palabras claves: Identidad biométrica, interoperabilidad, reconocimiento facial, aprendizaje profundo, aprendizaje de transferencia, clasificación de imágenes, BigData, contenedores, sociedad digital, economía digital.

ABSTRACT – KEY WORDS

The field of facial recognition has experienced constant advances in the last two decades, until reaching the accurate real-time recognition systems available today in cell phones and social networks. Currently, artificial intelligence has evolved so much that now exists even as a service Amazon Rekognition[4], as well as Apple's Face ID[5].

This thesis proposes a unique biometric identity solution, an important solution in our country, since it will allow us to identify a person in digital environments, digitally verifying the person who claims to be, and with this identification to perform procedures online, without the need to go personally to a government office, thus enabling a society and a digital economy and leaving a digital traceability.

The identification of the unique biometric identity is an important solution for a citizen who lives anywhere in the country, can perform a procedure online and remotely, using video cameras with what will help the government official to attend online and remotely, saving costs and time in attention using the digital one stop.

The Peruvian state has raised a law for the single digital identity, and aims to enable online authentication of the identity of Peruvian and foreign people when they need to access digital services provided by public entities, this law was published in September 2018, with the legislative decree No. 1412, in Chapter II digital identity, which drives the implementation of the national platform for identification and authentication of digital identity.

The presented thesis has a proposed scope from the biometric registration, and the generation of the digital identity credential, as well as the mechanisms of authentication and digital authorization, using biometric recognition and digital signatures, to enable a digital interoperability [5] using X-Road[6].

This thesis proposes a single biometric identity solution using face identification libraries with OpenCV [1], as well as for face recognition using deep learning techniques, and transfer learning.

[4] Amazon Rekognition, Rekognition's API and service will identify objects, people, text, scenes and activities, as well as detect inappropriate content. https://aws.amazon.com/es/rekognition/.
[5] Face ID, authentication through facial recognition, Apple https://support.apple.com/en-us/HT208109.
[6] X-Road, Interoperability Platform, https://x-road.global/architecture

Comparisons were made with the different architectures VGG16[2], VGG19[2] and RESTNET50 [3], generating a training accuracy greater than 98%, working with 10 classes and each class with 300 images of 224x224 pixels per input image for each class, supported by TensorFlow, Keras, Sklearn libraries, and validating the performance of the models using the confusion matrix, as well as for their storage of the identification and authentication traceability in BigData, and the use of containers (kubernetes/docker) to have a national and global scaling level. Each of these libraries and the proposed solution are 100% free software.

Keywords: Biometric Identity, Interoperability, Face Recognition, Deep Learning, Transfer Learning, Image Classification, BigData, Containers, Digital Society, Digital Economy, digital one stop.

TABLA DE CONTENIDOS

Página

Acrónimos y Abreviaturas 12

Lista de Figuras 15

Lista de Tablas 17

CAPÍTULO 1 18

PLANTEAMIENTO DEL PROBLEMA 18

1.1 Antecedentes 18

1.2 Delimitaciones y Definición del Problema 22

1.2.1 Delimitaciones 22

1.2.2 Definición del Problema 22

1.3 Formulación del Problema 23

1.4 Objetivo de la Investigación 25

1.4.1 Objetivo General 25

1.4.2 Objetivos Específicos 25

1.5 Justificación 26

1.6 Variable e Indicadores 26

1.7 Viabilidad de la Investigación 28

1.7.1 Viabilidad Técnica 28

1.7.2 Viabilidad Operativa 28

1.7.3 Viabilidad Económica 28

1.8 Contribución 28

1.9 Limitaciones de la Investigación 29

1.10 Tipo de Investigación 30

1.11 Nivel de Investigación 30

1.12 Método de la Investigación 30

1.13 Diseño de la Investigación 30

1.14 Técnicas e Instrumentos de Recolección de Información 31

1.14.1 Técnicas 31

1.14.2 Instrumentos 31

1.15 Cobertura de Estudio 31

1.15.1 Universo 31

1.15.2 Muestra 31
1.16 Estructura de la Tesis 31
CAPÍTULO 2 33
ESTADO DEL ARTE 33
2.1 Identificación Biométrica 33
2.2 Redes Neuronales Convolucionales (CNN) 35
CAPÍTULO 3 38
MARCO TEÓRICO 38
3.1 Detección de Rostros usando Haar Cascades 38
3.2 Deep Learning y TensorFlow 41
3.2.1 Deep Learning 41
3.2.2 Estructura de las Redes Neuronales Convolucionales 41
3.2.3 La Capa de Convolución 42
3.2.4 Capa de Reducción o Pooling 45
3.2.5 Capa clasificadora totalmente conectada 46
3.2.6 SoftMax 47
3.2.7 Dropout 47
3.2.8 Arquitectura CNN 48
3.2.9 Funciones de Activación 49
3.2.10 Tensorflow y Tensores 49
3.2.11 Big Data 54
3.2.12 Matriz de Confusión 57
3.2.13 Interoperabilidad 58
CAPÍTULO 4 61
DESARROLLO 61
4.1 Diseño de la Propuesta 61
4.2 Recolección y Sistematización de Recolección de Conjunto de Imágenes 62
4.2.1 Registro de identidad Biométrica Única del ciudadano 64
4.2.1.1 Narrativa 65
4.2.1.2 Proceso: Registro de identidad Biométrica Única del ciudadano 66
4.3 Entrenamiento de Redes Neuronales Convolucionales y su Evaluación 67
4.3.1 Autenticación e identificación Biométrica 69
4.3.1.1 Narrativa 70

4.3.1.2 Proceso: Autenticación e Identificación Biométrica 70
4.4 Aspectos de Infraestructura 71
4.5 Estrategia de implementación 71
4.5.1 Cadena de valor End-to-End de la solución Identificación Biométrica Única 71
4.6 Diseño de implementación 73
4.7 Diseño de Arquitectura 75
4.7.1 Arquitectura Semántica 75
4.7.2 Arquitectura Tecnológica 80
4.7.3 Arquitectura de Solución 83
4.8 Comunicaciones Segura 87
4.9 Escalamiento horizontal 87
CAPITULO 5 89
RESULTADOS, PRUEBAS Y EVALUACIÓN 89
5.1 Módulos del Sistema 89
5.2 Pruebas y Resultados 95
5.3 Evaluación de Resultados 99
5.1.1 Evaluación VGG16 99
5.1.2 Evaluación VGG19 101
5.1.3 Evaluación RESTNET50 102
CONCLUSIONES Y RECOMENDACIONES 104
CONCLUSIONES 104
RECOMENDACIONES 106
REFERENCIAS BIBLIOGRÁFICAS 107
LINK DE INTERNET 115
ANEXO A: Configuración del Modelo VGG16+3 116

Acrónimos y Abreviaturas

Abreviaturas	Descripción
Convolutional Neural Network (CNN)	Red neuronal convolucional, son un tipo de redes neuronales artificiales donde las "neuronas" corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria (V1) de un cerebro biológico [22].
VGG16[7]	Visual Geometry Group. Topología y Arquitectura de red neuronal convolucional de 13 capas convolucionales y 3 densas.
VGG19[8]	Visual Geometry Group. Topología y Arquitectura de red neuronal convolucional de 16 capas convolucionales y 3 densas.
RestNet50[9]	Deep Residual Learning for Image Recognition, Topología y Arquitectura de red neuronal convolucional de 152 capas.
Transfer Learning[10]	El aprendizaje por transferencia es un método de aprendizaje automático en el que un modelo desarrollado para una tarea se reutiliza como punto de partida para un modelo en una segunda tarea.
Deep Learning	El deep learning es un tipo de machine learning que entrena a una computadora para que realice tareas como las que hacemos los seres humanos, como el reconocimiento del habla, la identificación de imágenes o hacer predicciones de datos estadísticos.
SVM: Support Vector Machine	las máquinas de vectores de soporte (SVM), también redes de vectores de soporte [59]) son modelos de aprendizaje supervisado con algoritmos de aprendizaje asociados que analizan datos para el análisis de clasificación y regresión.

[7] https://www.robots.ox.ac.uk/~vgg/
[8] https://www.robots.ox.ac.uk/~vgg/
[9] https://arxiv.org/abs/1512.03385
[10] https://www.viewnext.com/transfer-learning-y-redes-convolucionales/

	los vectores de entrada se mapean de forma no lineal en un espacio de características de muy alta dimensión. En este espacio de características se construye una superficie de decisión lineal[59].
Abandono (Dropout)	Es una técnica de regularización para reducir el sobreajuste en redes neuronales artificiales. Es una forma eficiente de realizar promedios de modelos con redes neuronales. El término dropout significa "abandonar" u omitir aleatoriamente neuronas durante el proceso de entrenamiento de una red neuronal.
Matriz de confusión	En el campo de la inteligencia artificial y en especial en el problema de la clasificación estadística, una matriz de confusión es una herramienta que permite la visualización del desempeño de un algoritmo que se emplea en aprendizaje supervisado. Visualiza la distribución de los registros en términos de clases reales y clases pronosticadas. Esto indica la calidad del modelo actual. Un modelo puede contener dos o más clases pronosticadas.
Interoperabilidad	En un nivel muy básico, una interfaz de programación de aplicaciones (API), es una manera para que dos aplicaciones (sistemas) de computadora se hablen entre sí (compartan información, procesos y estándares) a través de una red utilizando un lenguaje común que ambas entienden (Jacobson et al., 2012).
Servidor Streaming	Un servidor Streaming es la infraestructura que da soporte a la transmisión de video o de audio en tiempo real,
Autoridad de certificación (CA[11])	Señalan a una entidad de confianza, responsable de emitir y revocar los certificados, utilizando en ellos la firma electrónica, para lo cual se emplea la criptografía de clave pública.

[11] Creación de un certificado de CA en SSL, https://www.ibm.com/docs/es/was/9.0.5?topic=configuration-creating-ca-certificate-in-ssl

Solicitud de firma de certificado (CSR .csr)	Es un mensaje enviado por un solicitante a una autoridad de registro de la infraestructura de clave pública para solicitar un certificado de identidad digital.
Private Key (.key)	El archivo de clave privada utilizado por la CA junto con CSR para firmar y generar los certificados.
Certificado (.crt)	El certificado generado por la CA.
Certificado auto firmado SSL/TLS	Un certificado auto firmado es un certificado SSL que no ha sido validado por una Autoridad de certificación (CA). El nivel de cifrado puede ser el mismo que el de cualquier otro certificado, pero debido a que no está validado por una CA, el navegador mostrará una advertencia cuando se visite el sitio.
OTT (Over-the-top media services)	En telecomunicaciones, un servicio de libre transmisión o servicio OTT consiste en la transmisión de audio, vídeo y otros contenidos a través de internet sin la implicación de los operadores tradicionales en el control o la distribución del contenido. OTT, a veces también se llama streaming de TV, representa cualquier contenido que se distribuya directamente a los espectadores a través de un servicio de transmisión de video a través de Internet, sin usar la caja de cable tradicional, y que regularmente se ve en una televisión. (https://nexttv.github.io/)
ImageNet	Es una base de datos de imágenes organizada según la jerarquía de WordNet (actualmente solo los sustantivos), en la que cada nodo de la jerarquía está representado por cientos y miles de imágenes. El proyecto ha sido fundamental en el avance de la visión por computadora y la investigación de aprendizaje profundo. Los datos están disponibles de forma gratuita para los investigadores para uso no comercial en https://www.image-net.org/.

Lista de Figuras

Figura 1, Arquitectura de la plataforma de Interoperabilidad Estonia (fuente X-ROAD) 19
Figura 2, Arquitectura X-Road Playground (fuente X-ROAD) 19
Figura 3, Alcance de la solución de comparación Biométrica (Biometric Exit, biometrics.cbp.gov) 21
Figura 4, Horas necesarias para completar un trámite, por país (fuente BID-2017) 23
Figura 5, Historia evolutiva de las CNN profundas 36
Figura 6, Evolución de las funciones de pérdida discriminativas 37
Figura 7, Detección de rostro y extracción rostro (fuente propia) 41
Figura 8, Inteligencia Artificial y Deep Learning (MIT- https://deeplearning.mit.edu) 41
Figura 9, Simple operación convolucional 43
Figura 10, la operación de convolución se indica con un asterisco (*), I*K 43
Figura 11, Operaciones en la Capa convolucional [56] 44
Figura 12, Filtro de convolución 45
Figura 13, Arquitectura de una red neuronal convolucional 48
Figura 14, Las funciones de activación en Tensorflow (fuente 6S191_MIT_DeepLearning_L1) 49
Figura 15, Estructura de una máquina y de un sistema distribuido [10] 50
Figura 16, Visualización gráfica en TensorBoard de un modelo de red neuronal convolucional [10], ... 51
Figura 17, Matriz de confusión 58
Figura 18, Evaluación de desempeño de modelos entrenados[65] 58
Figura 19, Modelo de interoperabilidad de gobierno electrónico [5] 60
Figura 20, Diseño de la propuesta para acceder a realizar un trámite de forma remota 62
Figura 21, registro y extracción de la foto del rostro del ciudadano (fuente propia) 63
Figura 22, Ejemplo de una muestra de 50 de las 300 imágenes recolectadas (224x224 resolución por imagen - fuente propia) 64
Figura 23, Registro de identidad Biométrica Única del ciudadano 66
Figura 24, Modelo de entrenamiento de la CNN (fuente propia) 67
Figura 25, Modelo de entrenamiento de las redes neuronales convolucionales (fuente propia) 68
Figura 26, Autenticación biométrica del Ciudadano 70
Figura 27, Evaluación del entrenamiento y reconocimiento facial (fuente propia) 71
Figura 28, Cadena de Valor End to End para la Identificación Biométrica Única (fuente propia) 72
Figura 29, Factores de Influencia en la Estrategia de Interoperabilidad (fuente propia) 72
Figura 30, Diseño de implementación (fuente propia) 73
Figura 31, Identidad Biométrica 76
Figura 32, Modelo conceptual de la arquitectura tecnológica de solución (fuente propia) 81
Figura 33, Modelo de enrolamiento biométrico 84
Figura 34, Arquitectura de Modelos VGG16, VGG19, RestNet50 84
Figura 35, Arquitectura VGG19 85
Figura 36, Evaluación Biométrica 85
Figura 37, Arquitectura Conceptual de Identificación Biométricos Única (Fuente Propia) 86
Figura 38, Ejemplo del módulo listar rostros registrados(fuente propia) 89
Figura 39, Módulos del sistema de identificación biométrica única 90
Figura 40, Recorrido horizontal de las etapas para el cálculo del haar [13]. 92
Figura 41, Operación para el cálculo de haar, la existencia de contorno 92

Figura 42, Enrolamiento, identificación y extracción de rostros utilizando OpenCV 93
Figura 43, la imagen integral para calcular el haar 93
Figura 44, lista de las 9 clases utilizadas para el entrenamiento de los diferentes modelos 94
Figura 45, Mapa de características por bloques de la red neuronal convolucional entrenada 94
Figura 46, Visualización de características del primer bloque utilizando VGG16 (block1_conv1 4x16).. 95
Figura 47, Visualización de características del segundo bloque utilizando VGG16 (block2_conv1 8x16)95
Figura 48, resultado y evaluación de los modelos VGG16, VGG19, RESTNET50 y VGG16+3 (fuente propia) 96
Figura 49, Evaluación del desempeño de modelos entrenados VGG16, VGG19 y RESTNET50 97
Figura 50, Evaluación de la red neuronal entrenada 98
Figura 51, Logs generado en el proceso de reconocimiento facial 98
Figura 52, trazabilidad de autenticación biométrica (fuente propia) 99
Figura 53, Evaluación matriz de confusión para VGG16+3 101
Figura 54, Evaluación matriz de confusión para VGG19 102
Figura 55, Evaluación matriz de confusión para RESTNET50 103

Lista de Tablas

Tabla 1, Variables e indicadores 27
Tabla 2, Ejemplo de operación de convolución 44
Tabla 3, Modelos disponibles KERAS, número de parámetros, precisión y tiempo (fuente keras.io) 52
Tabla 4, Vista del modelo de datos Cassandra [10] 57
Tabla 5, Conjunto de datos usados en la propuesta (fuente propia) 63
Tabla 6, Registro de identidad Biométrica Única del Ciudadano 64
Tabla 7, Configuración de las redes neuronales para su entrenamiento 68
Tabla 8, Autenticación de Identidad Biométrica Única del Ciudadano 69
Tabla 9, Configuración de infraestructura para realizar pruebas 71
Tabla 10, Diseño de implementación (fuente propia) 73
Tabla 11, Entidad de Identificación biométrica única (fuente propia) 75
Tabla 12, Entidad para la trazabilidad de autenticación biométrica única (fuente propia) 77
Tabla 13, Servicios utilizados para habilitar la identificación biométrica única (fuente propia) 78
Tabla 14, Generación de certificado SSL de seguridad para la plataforma de Identidad Biométrica 79
Tabla 15, Componentes tecnológicos utilizandos 81
Tabla 16, Lista de componentes implementados en la solución 90
Tabla 17, Tabla de comparación de resultados para los modelos VGG16, VGG19 y RESTNET50 96
Tabla 18, Tabla de evaluación del modelo VGG16+3 100
Tabla 19, Tabla de evaluación del modelo VGG19 102
Tabla 20, Tabla de evaluación del modelo RESTNET50 103

CAPÍTULO 1

PLANTEAMIENTO DEL PROBLEMA

1.1 Antecedentes

Francia aprobó un decreto 019-452 el 13 de mayo de 2019, autorizando la creación de un medio e identificación electrónica denominado Alicem (para 'Autenticación en ligne certifiée sur mobile', o Identificación Certificada en Línea en Teléfono Móvil). solución similar de autenticación como plantea la solución Apple ID[12], Alicem[1] una nueva aplicación gubernamental francesa, basada en software de reconocimiento facial, dará a los ciudadanos de Francia acceso a alrededor de 500 sitios web gubernamentales.

México aprobó y público las disposiciones generales y el reglamento de la firma electrónica avanzada 2016, adoptando controles de seguridad como una autoridad certificadora, generación de claves criptográficas, protocolo seguro https, así como OCSP un servicio de validación en tiempo real del estado actual de un certificado, y funciona como identificador único para ingresar a trámites digitales en el caso de personas físicas y morales en el portal de gob.mx, y tiene los mismos efectos jurídicos que la firma autógrafa (firma de identificación).

Perú aprobó la actualización de la Ley N° 27444 el 2 de febrero de 2019, Ley del Procedimiento Administrativo General, indica que las entidades deben "Proporcionar directamente los datos e información que posean, sea cual fuere su naturaleza jurídica o posición institucional, a través de cualquier medio, sin más limitación que la establecida por la Constitución o la ley, para lo cual se propenderá a la interconexión de equipos de procesamiento electrónico de información, u otros medios similares[13]", pero para poder interoperar e intercambiar información deben existir principios que regulen este intercambio de información, como son la no repudiación, integridad de la información, estandarización, seguridad de la información con terceros, garantías de entrega y

[12] Noviembre 2017, RESEARCH AREA COMPUTER VISION, An On-device Deep Neural Network for Face Detection https://machinelearning.apple.com/research/face-detection
[13] Artículo 87.2.2 del TUO de la Ley N° 27444, aprobado con el Decreto Supremo N° 004-2019-JUS.

sobre todo una identidad digital única aprobada en el decreto legislativo nº 1412, en el capítulo II identidad digital.

Estonia, implementó en el 2001 una plataforma de interoperabilidad denominada **X-Road**[14] es una capa de intercambio de datos distribuidos entre sistemas de información que proporciona una forma **estandarizada** y segura de producir y consumir servicios, garantizando la confidencialidad, la integridad y la interoperabilidad entre las partes que intercambian datos.

Técnicamente, el ecosistema de X-Road consta de servicios centrales, servidores de seguridad, sistemas de información, TSA (s) y CA (s).

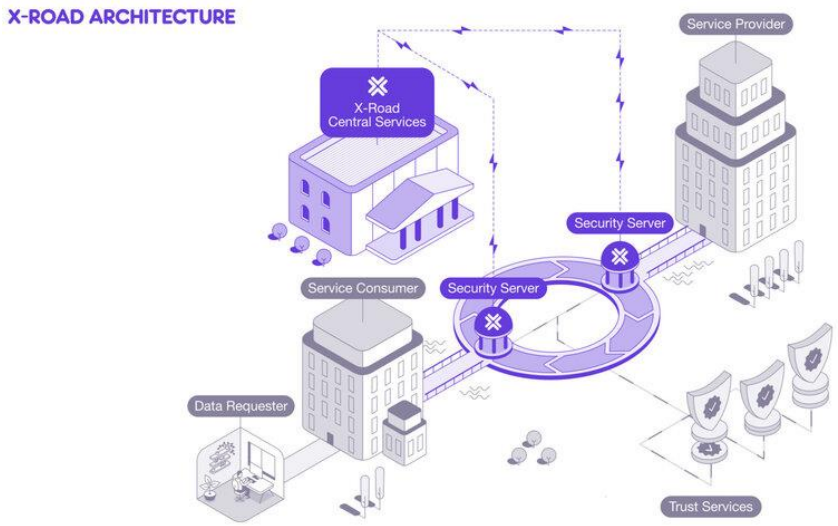


Figura 1, Arquitectura de la plataforma de Interoperabilidad Estonia (fuente X-ROAD)

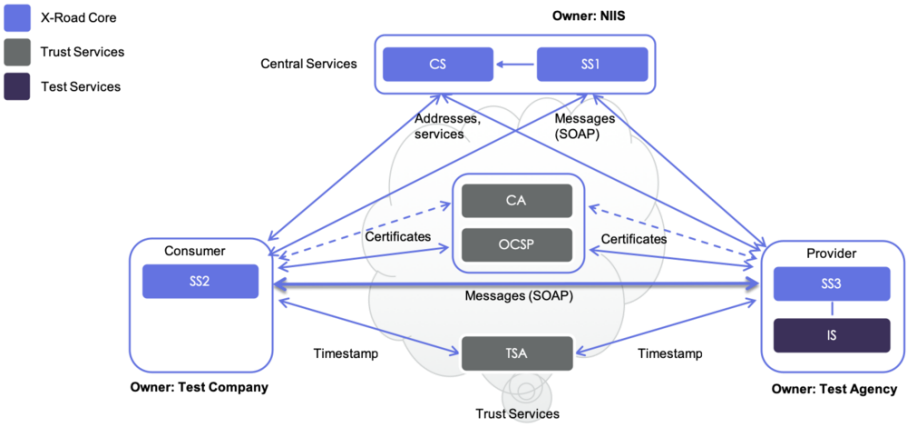


Figura 2, Arquitectura X-Road Playground (fuente X-ROAD)

[14] X-ROAD, https://x-road.global/x-road-technology-overview

X-Road Playground es un entorno X-Road completamente funcional que consta de:

- Servidor central (CS): Central Server administra la base de datos de los miembros de X-Road y los servidores de seguridad. Además, el servidor central contiene la política de seguridad de la instancia de X-Road. La política de seguridad consta de los siguientes elementos:
    - Lista de autoridades de certificación de confianza,
    - Lista de autoridades confiables de sellado de tiempo,
    - Parámetros ajustables, como la vida útil máxima permitida de una respuesta OCSP.
- Servidores de seguridad (SS): Security Server encapsula los aspectos de seguridad de la infraestructura X-Road: administrar claves para firma y autenticación, enviar mensajes a través de un canal seguro, crear el valor de prueba para los mensajes con firmas digitales, sellado de tiempo y registro. Para el cliente de servicios y el sistema de información del proveedor de servicios, Security Server ofrece un protocolo de mensajes para SOAP y un protocolo de mensajes para REST.
- Servicios de prueba (IS): El sistema de información (IS) utiliza y / o proporciona servicios a través de X-Road.
- Servicios de confianza (CA, OCSP, TSA):
    - TSA: La autoridad de sellado de tiempo emite sellos de tiempo que certifican la existencia de elementos de datos en un momento determinado.
    - CA: La autoridad de certificación (CA) emite certificados a los servidores de seguridad (certificados de autenticación) y a las organizaciones miembro de X-Road (certificados de firma). Todos los certificados se almacenan en los servidores de seguridad. La CA debe poder procesar solicitudes de firma de certificados de conformidad con PKCS10.
    - OCSP: La CA debe distribuir la información de validez del certificado a través del protocolo OCSP. Los Security Servers almacenan en caché las respuestas OCSP

para reducir la carga en el servicio OCSP y aumentar la disponibilidad. La carga del servicio OCSP depende del número de certificados emitidos.

Estados Unidos, La tecnología de comparación facial se implementará en 198 aeropuertos, que incluyen 14 ubicaciones internacionales. Biometric Exit [15] ya empezó implementando en 29 aeropuertos. Hay 11 puertos marítimos donde lo utilizan las líneas de cruceros. A lo largo de las fronteras terrestres del sur de estados unidos, miles de viajeros que van y vienen del trabajo de forma diaria, encuentran la tecnología como una forma de hacer sus viajes diarios más rápidos y eficientes y su alcance es aéreo, terrestre y marítimo.

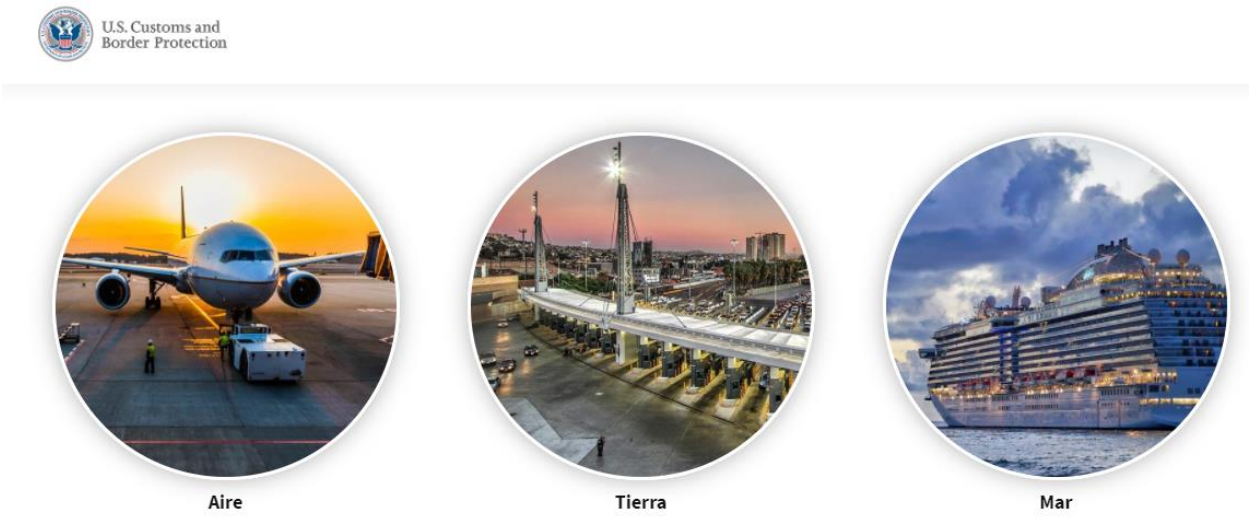


Figura 3, Alcance de la solución de comparación Biométrica (Biometric Exit, biometrics.cbp.gov)

Actualmente la **PIDE**[16] (plataforma nacional de interoperabilidad del estado peruano) permite la implementación de servicios públicos en línea que facilita la simplificación administrativa y los procesos de negocio de las instituciones, sin embargo esta plataforma solo es publicación de servicios a través de servicios web para que otras instituciones puedan consumir estos servicios.

[15] Introducing Biometric Facial Comparison, https://biometrics.cbp.gov/#How-it-works
[16] Plataforma Nacional de Interoperabilidad, https://www.gob.pe/741-plataforma-de-interoperabilidad-del-estado

1.2 Delimitaciones y Definición del Problema

1.2.1 Delimitaciones

El alcance de la propuesta es identificar biométricamente a una persona en un entorno digital utilizando técnicas de visión por computadora, aprendizaje profundo y aprendizaje por transferencia.

La propuesta se centrará en una aplicación web, que contará con de registro, entrenamiento, verificación y autenticación utilizando credenciales criptográficas y características biométricas faciales.

La tesis es parte de una propuesta a la implementación de la identidad biométrica única para el estado peruano (IBU).

1.2.2 Definición del Problema

La falta una identidad biometría única limita la interoperabilidad entre las instituciones y la falta de una ventanilla única nacional para el estado, genera que cada institución desarrolle su propio portal de trámites y servicios, así como su propia gestión de identidades.

Los ciudadanos para autenticarse a un portal de una institución A requieren unas credenciales y otras credenciales diferentes cuando acceden a otra institución B, generando un problema de estandarización, de trazabilidad y de interoperabilidad. Según el último análisis y encuesta de las Naciones Unidas[17], Perú se encuentra en el lugar 77, con un índice de desarrollo en gobierno electrónico con un 64.61%.

Sin embargo, es importante por parte del gobierno central establecer normativas que ayuden en base a procedimientos y políticas a poner condiciones para que el ciudadano no pierda tanto tiempo y costo haciendo un trámite, y que el gobierno reduzca considerablemente su gasto operativo para gestionar y administrar un trámite de forma eficiente, y así poder optimizar los recursos del estado y ahorrar tiempo y dinero a los ciudadanos.

[17] Reporte anual Naciones unidas 2018, E-Government Development Index https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2018-Survey/E-Government%20Survey%202018_FINAL%20for%20web.pdf

1.3 Formulación del Problema

Primarias:

- ✓ ¿Es posible que con la implementación de una solución de identificación biométrica única pueda habilitar la interoperabilidad global en el Perú?

  Importantes avances realizados en el aeropuerto de Atlanta a partir del 2017, en Aduanas y protección de fronteras de estados unidos, demuestran el éxito de estas tecnologías biométricas.

  La tecnología de identificación biométrica que se usa en este caso consiste en una cámara que compara los rostros que escanea de los viajeros con los que tiene registrados en una base de datos verificada por la Oficina de Aduanas y Protección Fronteriza (CBP) de Estados Unidos, ahorrando tiempo y gestión administrativa.

- ✓ ¿Cuánto tiempo demora un ciudadano para realizar un trámite de forma presencial en la administración pública?

  El siguiente grafico muestra una estadística del tiempo que demora un ciudadano en realizar un trámite en una organización gubernamental, en Perú se tarda aproximadamente 8,6 horas, siendo un promedio en Latinoamérica de 5,4 horas, tardando el Perú, 3,2 horas encima del promedio, como se visualiza en la siguiente figura.

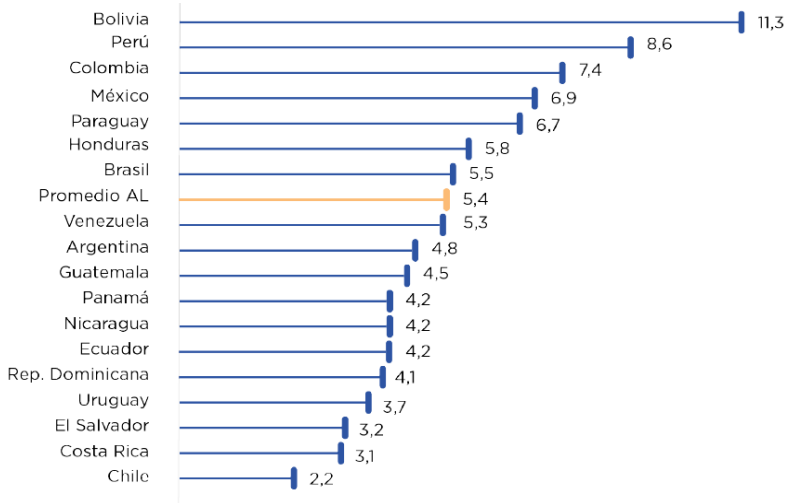


Figura 4, Horas necesarias para completar un trámite, por país (fuente BID-2017)

Encuesta realizada a hogares de 18 países de ALC, con un total de 20,200 personas encuestadas.

- ✓ ¿Cuánto le cuesta al estado peruanos atender en ventanilla un trámite de forma presencial a un ciudadano?

  En promedio le cuesta al estado 207 dólares, según la investigación realizado por el BID en el 2017.

- ✓ ¿Cómo se mide el tiempo para realizar un trámite?

  En promedio, los ciudadanos de la región se demoran 5,4 horas en hacer un trámite, según el estudio del BID: "El fin del trámite eterno[18]", en **Perú realizar un trámite se demora 8.6** horas aproximadamente.

  Se debe considerar desde que un ciudadano sale de su domicilio u oficina, hasta que se termine el trámite, y este en el lugar de origen.

  Se considera los siguientes factores:

  1. Leer un aviso del trámite
  2. Solicitud de la cita en línea
  3. Llenar un formulario
  4. Pago del trámite en los bancos (horarios y días con mayor afluencia)
  5. Cita en la oficina gubernamental para realizar el trámite
  6. Tiempo de espera de resolución de trámite según sea el caso
  7. Hacer fila y esperar en ventanilla
  8. Factores externos (tráfico, huelgas/paros, construcciones, accidentes, otros)

- ✓ ¿Cuánto contamina un ciudadano cuando va realizar un trámite a una institución pública de forma presencial?

  Es importante considerar si el ciudadano va en su propio transporte o en transporte público, o en casos extremos tiene que viajar a Lima y realizar sus trámites ejemplo los trámites de relaciones exteriores.

[18] Wait No More: Citizens, Red Tape, and Digital Government, https://publications.iadb.org/en/wait-no-more-citizens-red-tape-and-digital-government.

Secundarias:

- ✓ ¿Existe la normativa regulatoria para poder acceder a trámites y servicios identificando a una persona por su biometría facial utilizando video cámara y por internet?

  No, pero en Perú se aprobó el documento nacional de identidad digital DNId[19], que emitirá la RENIE en dispositivos digitales, que acreditará a la persona en entornos presenciales y no presenciales.

- ✓ El plástico utilizado en las identificaciones de los ciudadanos, ¿Qué tanto contaminan al medio ambiente?
  ONU [20], el 80% de basura marina está compuesto por plástico.
  Un abolsa de plástico dura aproximadamente 150 años en descomponerse, la Ley 30884, regula el plástico de un solo uso y los recipientes o envases descartables.

1.4 Objetivo de la Investigación

1.4.1 Objetivo General

Crear una solución de identidad biométrica única para habilitar la interoperabilidad global peruana utilizando técnicas de Deep Learning (redes neuronales convolucionales VGG16, VGG19 y RESTNET50) y una base de datos NoSQL columnar cassandra (BigData) para registrar la trazabilidad de eventos de autenticación biométrica.

1.4.2 Objetivos Específicos

1. Implementar la identidad biométrica única utilizando redes neuronales convolucionales VGG16, VGG19 y RESTNET50, y Transfer Learning como mecanismo de almacenamiento de conocimiento adquirido generado por las redes neuronales convolucionales propuestas. Y para registrar su trazabilidad de autenticación biométrica, cassandra (Base de datos columnar NoSQL).

[19] Identidad Digital Perú, https://idperu.reniec.gob.pe/site/
[20] Cuanto tiempo tardan los plásticos en descomponerse, https://www.eltiempo.com/historias-el-tiempo/cuanto-tiempo-tardan-los-plasticos-en-descomponerse-342568

2. Implementar mecanismo de extracción de características biométricas en particular rostros de una imagen y/o cámara web, utilizando técnicas de visión artificial por computadora, en particular el algoritmo de Haar Cascade implementadas en las librerías de OpenCV desarrolladas por Intel.
3. Elaborar una arquitectura de solución para soportar alta disponibilidad de servicios biométricos, de accesibilidad y de seguridad utilizando X-ROAD.
4. Evaluar los algoritmos de redes neuronales convolucionales VGG16, VGG19, y RESTNET50 propuestos, utilizando la matriz de confusión.

1.5 Justificación

La investigación propuesta tiene una oportunidad evidente por tratarse de un tema de suma actualidad, la Política Nacional de Gobierno Electrónico 2013-2017, aprobada mediante el Decreto Supremo N° 081- 2013-PCM, establece como uno de sus objetivos "Acercar el Estado a los ciudadanos, de manera articulada, a través de las tecnologías de la información que aseguren el acceso oportuno e inclusivo a la información y participación ciudadana como medio para contribuir a la gobernabilidad, transparencia y lucha contra la corrupción en la gestión del Estado" [21], así como el Decreto Legislativo N° 1412, publicado el 13 de setiembre de 2018, cuyo objetivo es permitir la autenticación en línea de la identidad de las personas peruanas y extranjeras cuando necesiten acceder a los servicios digitales brindados por las entidades públicas. También, proporciona el servicio de autenticación de la identidad digital a todas las instituciones públicas cuando desarrollen un servicio digital.

1.6 Variable e Indicadores

La siguiente tabla muestra las variables independientes, así como las dependientes y sus respectivos indicadores.

[21] El peruano / viernes 23 de marzo 2018, normas legales. https://sgp.pcm.gob.pe/wp-content/uploads/2018/05/Decreto-Supremo-N%C2%BA-031-2018-PCM.pdf

Tabla 1, Variables e indicadores

VARIABLES		**INDICADORES**
Independiente	Software de identificación biométrica	Lenguaje de Programación
		Base de datos
		Arquitectura
		Contenedores
		Seguridad
Dependiente	Autenticación Única del ciudadano	Autenticación
		Autorización
		Federación
		Detección de Cara
		Extracción de característica Biométrica
		Reconocimiento
		Gestión de Contenedores y Big Data
		Comunicación HTTPS
	Distribución	Demanda Nacional
		Plan de Marketing

1.7 Viabilidad de la Investigación

1.7.1 Viabilidad Técnica

1. El proyecto reúne características y condiciones que aseguran el cumplimiento de los objetivos y mediciones técnicas de precisión evaluando las tres técnicas VGG16, VGG19 y RESTNET50.
2. Gracias a la existencia de una referencia técnica y de interoperabilidad en el gobierno de Estonia[22] es factible su implementación técnica en el gobierno del Perú.
3. Todos los componentes y programas de software utilizados es 100% software libre.

1.7.2 Viabilidad Operativa

1. La solución propuesta tiene un acceso biométrico único, lo cual simplifica la accesibilidad con medios electrónicos y físicos, así como la utilización del internet.
2. La solución propuesta crea condiciones de polarización a través de la interoperabilidad.
3. La solución propuesta busca responder a la estrategia digital nacional del Perú.

1.7.3 Viabilidad Económica

1. El costo del manteamiento en el tiempo de la solución está dirigido directamente a recursos humanos, personal especializado en tecnologías de información, el gobierno peruano no pagaría ningún tipo de licenciamiento anual ya que el software utilizado es de libre distribución.
2. El impacto de carácter social esperado es considerable en cuanto al costo, y tiempo que un ciudadano y una institución pública/ privada resuelven el trámite.
3. Dentro del grupo de beneficiados (ciudadanos, instituciones privadas, instituciones públicas), son de carácter social que mejoran su calidad de vida.

1.8 Contribución

1. El estado peruano deberá implementar una ventanilla única digital de atención para gestionar trámites y servicios, donde el ciudadano podrá acceder y realizar sus

[22] Identidad Digital Emitida por el Estado, https://e-estonia.com/solutions/e-identity/

trámites, servicios y pagos en línea, evitando ir físicamente a una oficina de gobierno, ahorrando costos el ciudadano como la administración pública.

2. Contar con una arquitectura de solución de identificación biométrica única, utilizando mecanismos seguros de acceso oportuno a la información y una participación ciudadana, servirá como medio para aportar a la gobernabilidad y transparencia de la gestión del Estado Peruano, esto permitirá acelerar y habilitar una sociedad digital.[23]
3. Las posibilidades que generará esta solución en una institución de gobierno, y para el ciudadano son los siguientes:
    a. Ahorrar costos económicos y de tiempo de atención a las instituciones gubernamentales para los ciudadanos.
    b. Resolver trámites de forma remota, utilizando identificación biométrica, evitando el traslado a los ciudadanos y garantizando así su identidad y su trazabilidad.
    c. Habilitará la ventanilla única digital para que los ciudadanos puedan realizar trámites y servicios administrativos, desde sus oficinas, domicilios con medios digitales y de forma ágiles, simples, con mecanismos de autoservicio y seguros.
    d. Impactará de forma directa a la sociedad, acelerando una cultura y una sociedad digital.

1.9 Limitaciones de la Investigación

1. No está integrado a un trámite, ni a un servicio del estado peruano, ni de organizaciones privadas.
2. No está habilitado el mecanismo OTT, ni el servidor de streaming para la comunicación remota a través de una cámara web.
3. No está habilitado el procesamiento de entrenamiento, ni el reconocimiento de la red neuronal sobre GPU's, solo está configurado el procesamiento para CPU's.
4. El modelo no está entrenado con 30 millones de ciudadano/clases del Perú, el modelo de entrenamiento solo está entrenado con una muestra de 9 diferentes clases, y cada

[23] El peruano / viernes 23 marzo 2018, normas legales. (Plan Bicentenario: El Perú hacia el 2021, aprobado por Decreto Supremo N° 054-2011-PCM).

clase con 300 imágenes con una resolución de 224x224 de entrada por imagen, siendo un total de 2700 imágenes, utilizando el 80% para el entrenamiento y el 20% para la evaluación de la red neuronal convolucional [58].

1.10 Tipo de Investigación

1. Aplicada (No Experimental y Descriptivo)
   Crear una solución de identidad biométrica única.

1.11 Nivel de Investigación

Exploratorio y Explicativo descriptivo.

1.12 Método de la Investigación

Para detección de rostros, se utilizará técnicas de visión por computadora utilizando las bibliotecas de OpenCV [1].

Para el entrenamiento facial se utilizaron tres métodos, VGG16[2], VGG19[2] y RESTNET50[3], cada red entrenada con 10 iteraciones recomendadas por el estado del arte, así como para cada clase de imagen se utilizará 300 imágenes,

finalmente, los resultados serán medidos por los indicadores Precisión, exactitud utilizando la matriz de confusión como método de evaluación de resultados.

1.13 Diseño de la Investigación

Investigación transeccional (transversal) descriptivo, proporcionando la visión de la identidad biométrica única.

Para la etapa de recolección de datos se plantea realizar una aplicación que habilita la cámara frontal de la computadora, y esta captura las 300 imágenes de forma continua, para una muestra de 9 personas.

1.14 Técnicas e Instrumentos de Recolección de Información

1.14.1 Técnicas

- Entrevista
- Cuestionario
- Revisión bibliográfica
- Revisión documental

1.14.2 Instrumentos

- Entrevista personal
- Encuesta
- Textos
- Artículos
- Documentos de la organización

La información fue recolectada por una investigación realizada por el BID en el 2017 [60].

1.15 Cobertura de Estudio

1.15.1 Universo

Todas las organizaciones del Estado Peruano y Privados, donde un ciudadano tenga contacto y pueda realizar un trámite y/o servicio en línea.

1.15.2 Muestra

El muestreo fue de 2700 imágenes, y se seleccionarán 9 clases diferentes, y para cada clase se generaron 300 imágenes, tomando las cuotas para el entrenamiento del 80%(240 imágenes por clase, y con una resolución de 224x224 Píxeles por imagen, que son las entradas requeridas para los modelos VGG16, VGG19 y RESTNET50) y 20%(60 imágenes por clase, y una resolución de 224x224 Píxeles) para las evaluaciones.

1.16 Estructura de la Tesis

La presente tesis está organizada de la siguiente manera:

1. En el **Capítulo 1** se describe la formulación del presente documento, incluyendo los antecedentes, planteamiento del problema, los objetivos, las limitaciones, y las contribuciones.
2. En el **Capítulo 2** se hace una revisión del estado del arte de los mecanismos de identificación biométrica, y su evolución en el tiempo, así como su estado actual.
3. En el **Capítulo 3** se describe el marco teórico de la utilización de los algoritmos implementados utilizando la librería de OpenCV, en particular el Haar Cascade, conceptos de Deep Learning, Transfer Learning y redes neuronales convolucionales, tensores, contenedores, almacenamiento en BigData y sobre todo los habilitadores de la interoperabilidad conjuntamente con sus dimensiones.

4. En el **Capítulo 4** se describe el desarrollo de la solución, técnicas y algoritmos utilizados para la extracción biométrica del rostro (Haar Cascade), los procesos de negocio utilizados (enrolamiento biométrico, identificación biométrica), así como los mecanismos entrenamiento utilizando redes neuronales profundas como el VGG16, VGG19 y RESTNET50, y su respectiva evaluación de desempeño de los modelos de entrenamiento utilizando matriz de confusión.

5. En el **Capítulo 5** se describe los análisis de los resultados del desarrollo del Capítulo 4, y luego el análisis de las métricas de validación realizadas, así como la evaluación de los diferentes modelos RESTNET50, VGG16, VGG19 utilizando la matriz de confusión. Finalmente se describe la comparación de las métricas y modelos, en función de su rapidez, perdida y precisión

6. Conclusiones de la propuesta y recomendaciones sugeridas para su implementación en ambientes productivos con una escala nacional.

CAPÍTULO 2

ESTADO DEL ARTE

2.1 Identificación Biométrica

La tarea de reconocimiento facial descritos en [14], [25], [39], [42], [47], es diferente de las tareas de clasificación de objetos [2] genéricos debido a la particularidad de los rostros, tiene que lidiar con un número masivo de clases con pequeñas diferencias entre clases y grandes variaciones intrapersonales debido a diferentes poses, iluminaciones [56], expresiones, edades y oclusiones. Los métodos de aprendizaje profundo aprenden múltiples niveles de representaciones que corresponden a diferentes niveles de abstracción. Los niveles forman una jerarquía de conceptos, mostrando una fuerte invariancia a la pose facial, la iluminación y los cambios de expresión según la edad y su estado emocional.

Las tecnologías biométricas[61], [62], [64] y en particular el reconocimiento facial se están desarrollando rápidamente, y se con el tiempo estas tecnologías se convertirán en una extensión para mejorar y simplificar las actividades humanas y de muchas industrias como la robótica y la automatización, en la propuesta de la tienda de "Amazon Go & Amazon go Grocery" [26], [27] (teniendo en cuenta la computer vision[28], [29], Deep learning algorithms[30], [31], [32], [33],[34], [35], y el sensor fusión[36]), donde un usuario ingresa a la tienda y selecciona cualquier producto, y gracias al reconocimiento facial, clasificados en los siguientes trabajos como la detección parcial de rostros en base a SVM(super verctor machine) descrito por Upal Mahbub [17], [60], y micro-expresiones como lo menciona Kumar en [37], y seguimiento de caras en tiempo real utilizando técnicas de visión computacional como lo describe Qi Cao, Ruishan Liu en [38], así como el sistema de seguridad de acceso como lo plantea facefirst [39] , FaceLiveNet en autenticación [40] ,OpenFace en autenticación descritas por Giuseppe Garofalo [41], [42], [43], 19] y sensores que dan seguimiento en tiempo real, el usuario puede tomar cualquier producto, y el usuario puede salir de la tienda sin ningún mecanismo de bloqueo como una caja, una fila, servicio al cliente, y lo más importante y la contabilización de productos y el respectivo pago.

Alicem una solución que presenta Jerome Letier[44], significa "Authentification en Ligne Certifiée sur Mobile" (Autenticación en línea certificada en teléfonos móviles), solución similar de autenticación como plantea la solución Apple ID [45],[46], Alicem una nueva aplicación gubernamental francesa, basada en software de reconocimiento facial, dará a los ciudadanos de Francia acceso a alrededor de 500 sitios web gubernamentales, Jerome Letier (Director Digital del Ministerio del Interior,2020 - Francia), convencido de que la tecnología "Alicem" llevará a Francia a una nueva era digital, permitiendo a los ciudadanos acceder a los servicios a través de un sistema de alta seguridad sin que tengan que ir físicamente a una oficina gubernamental, es la propuesta de Lawmaker Didier desarrollada en agencia nacional de documentos en Francia et, al [44].

La tecnología de reconocimiento facial (FRT) hace posible comparar imágenes faciales digitales como la propuesta por Pawan Sinha y Ethan Meyers del MIT en [47], [48] para determinar si son de la misma persona. La comparación de imágenes obtenidas de cámaras de video (CCTV) con imágenes de bases de datos se denomina "tecnología de reconocimiento facial en vivo" como plantea Pete Fussey [49]. Los ejemplos de autoridades policiales nacionales de la UE (Unión Europea) que utilizan dicha tecnología son escasos, pero varios están probando su potencial. Por lo tanto, es importante analizar las implicaciones para los derechos fundamentales de confiar en FRT en vivo, centrándose en su uso para fines de trámites administrativos.

Las tecnologías de reconocimiento facial (FRT) propuesta por Nissen baum, Helen F. and Introna, Lucas D [50], [51] intentan verificar a las personas usando sus rostros. Los FRT detectan rostros, extraen rasgos distintivos de los rostros y comparan los rasgos con los acumulados en una base de datos. FRT quita el control de la propia identidad y le da ese control a un tercero como lo menciona Nissen baum [50], a menudo a las fuerzas del orden. El reconocimiento facial se puede utilizar como una identificación digital que permite al gobierno u otra entidad identificar a las personas en masa, sin consentimiento y en secreto. FRT amenaza la privacidad y las libertades civiles y las bases de datos requeridas para el reconocimiento facial a gran escala son vulnerables a las filtraciones de datos, el uso indebido y el avance de la misión, como se detalla en la política estatal de reconocimiento facial como plantea la "European Union Agency for Fundamental Rights" [20], [50], [52], y en las consideraciones de derechos fundamentales en el contexto de la aplicación de la ley, suscrita por la comunidad europea [20], [52].

China lleva años endureciendo las reglas para garantizar que todos los que usan internet lo hagan bajo su identidad real. A partir de ahora, quien contrate nuevos servicios de telefonía móvil en China tendrá que pasar por un escaneo facial, una nueva medida que causa polémica en el gigante asiático de acuerdo con BBC News Mundo [53].

Actualmente México cuenta con una identidad electrónica la e. Firma emitida por el SAT (servicio de administración tributaria), algunas instituciones la utilizan como un mecanismo de autenticación, sin embargo, la autenticación no está estandarizada en las 299 instituciones federales y sus más de 4 mil trámites de acuerdo con CONAMER (Catálogo Nacional de Regulaciones, Trámites y Servicios) [55].

Algunos factores importantes, es que la cantidad de funcionarios que trabajan en instituciones gubernamentales, no pueden abastecer a los más de 130 millones de mexicanos, los elevados costos de los tramites, los tiempos que un ciudadano tarda en realizar un trámite, la edad, tintes, barbas y los efectos del envejecimiento, la evaluación continua de las características biométricas como lo señala Philipp Terhorst, Daniel Fahrmann, Naser Damer,Florian Kirchbuchner and Arjan Kuijper [54], y que muchos trámites no están digitalizados aún de acuerdo a GOB.MX[55].

2.2 Redes Neuronales Convolucionales (CNN)

Las CNN [26], [57], [58] fueron diseñadas para tareas de reconocimiento de imágenes que originalmente se aplicaron al desafío del reconocimiento de dígitos escritos a mano. El objetivo de diseño básico de las CNN era crear una red en la que la neurona de la capa inicial de la red extraería características visuales locales, y las neuronas de las capas posteriores combinarían estas características para formar características de orden superior.

VGG-16 [2], [25], un modelo que alcanza una precisión de prueba del 92,7% entre los 5 primeros en ImageNet, que es un conjunto de datos de más de 14 millones de imágenes que pertenecen a 1000 clases. Fue propuesto por Karen Simonyan y Andrew Zisserman del Laboratorio del Grupo de Geometría Visual de la Universidad de Oxford en 2014.

VGG-19 [2], [25] es una red neuronal convolucional [57] que tiene 19 capas de profundidad y puede clasificar imágenes en 1000 categorías de objetos, como un teclado, un mouse y muchos animales. El modelo se entrenó con más de un millón de imágenes de la base de datos de ImageNet con una precisión del 92%VGG-19 es una red neuronal convolucional [57] que tiene 19 capas de

profundidad y puede clasificar imágenes en 1000 categorías de objetos, como un teclado, un mouse y muchos animales. El modelo se entrenó con más de un millón de imágenes de la base de datos de ImageNet con una precisión del 92%.

RESTNET [3], La red ha sido creada y familiarizada por Microsoft. Con una precisión del 96,4%, este modelo ganó el concurso ImageNet de 2016. Es conocido por su profundidad (hasta 152 capas) y la introducción de bloques residuales.

A continuación, se visualiza la evolución de las redes neuronales convolucionales en el tiempo, en la siguiente figura desde la primera definición de convolucion dada por la Neocognitron [23] en 1979 hasta Channel Bosted [24] en el 2018.

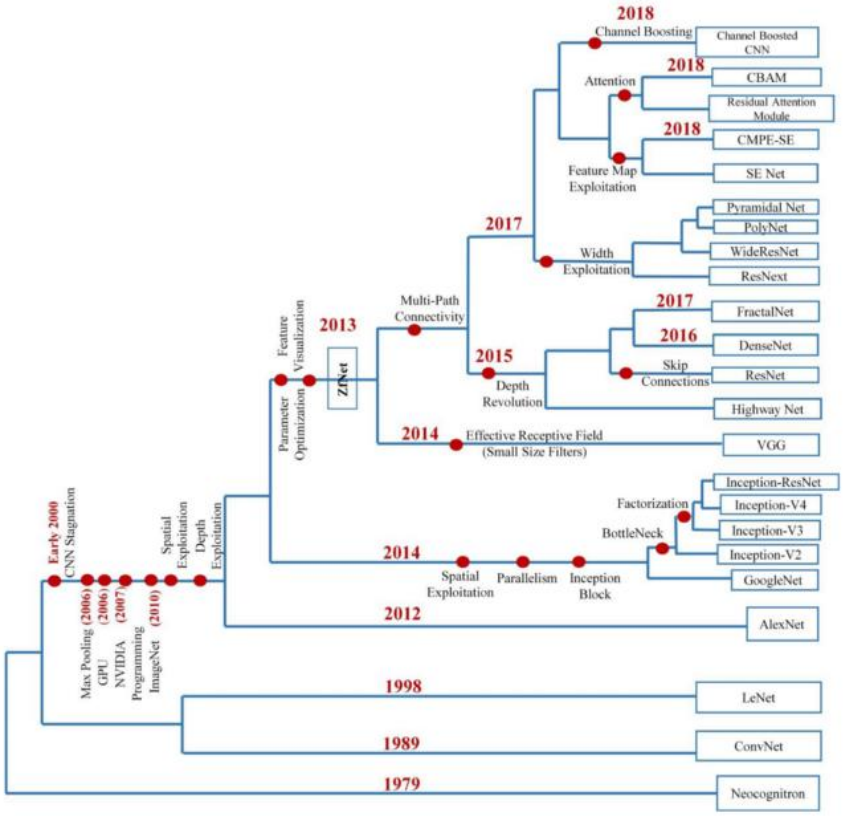


Figura 5, Historia evolutiva de las CNN profundas

La siguiente imagen muestra como en el tiempo ha ido evolucionando los algoritmos de perdida para la optimización en el entrenamiento de las redes neuronales convolucionales [25].

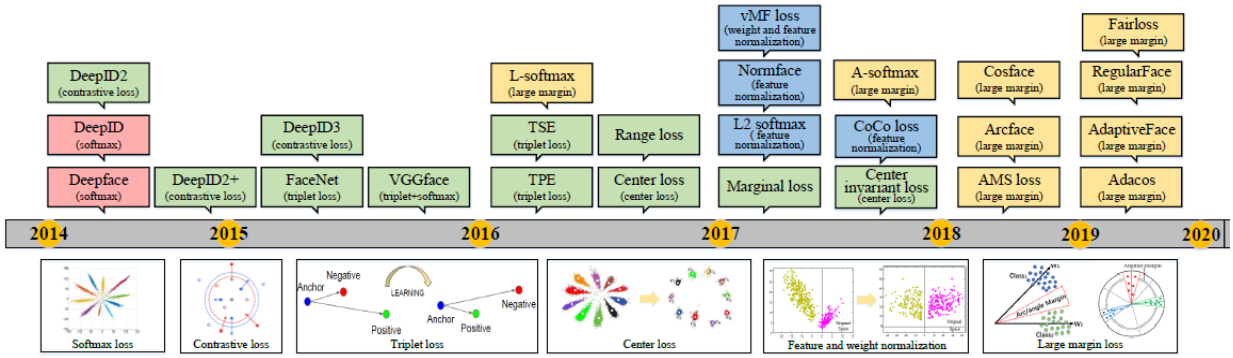


Figura 6, Evolución de las funciones de pérdida discriminativas

CAPÍTULO 3

MARCO TEÓRICO

En el siguiente Capítulo se describe la Revisión Histórica donde abarca teoría referente a Detección de Rostros y Deep Learning utilizando Tensorflow, BigData en particular Cassandra una base de datos Columnar, luego se abarca Aprendizaje Máquina y Profundo para entender cómo funcionan la Redes Neuronales Convolucionales y Arquitecturas de Detección de Rostros. Finalmente se describe el Estado del Arte que abarca las técnicas y conjuntos de datos que actualmente están siendo utilizados en la Literatura.

3.1 Detección de Rostros usando Haar Cascades

La Detección de objetos usando clasificadores en cascada basados en características de Haar Cascade es un método efectivo de detección de objetos propuesto por Paul Viola y Michael Jones en su artículo, "Detección rápida de objetos usando una cascada aumentada de características simples" en 2001. Es un enfoque basado en el aprendizaje automático donde La función de cascada se entrena a partir de muchas imágenes positivas y negativas. Luego se utiliza para detectar objetos en otras imágenes. Aquí trabajaremos con detección de rostros. Inicialmente, el algoritmo necesita muchas imágenes positivas (imágenes de caras) e imágenes negativas (imágenes sin caras) para entrenar al clasificador. Entonces necesitamos extraer características de él. Para esto, se utilizan las características de Haar que se muestran en la imagen de abajo. Son como nuestro núcleo convolucional. Cada característica es un valor único obtenido al restar la suma de píxeles debajo del rectángulo blanco de la suma de píxeles debajo del rectángulo negro.

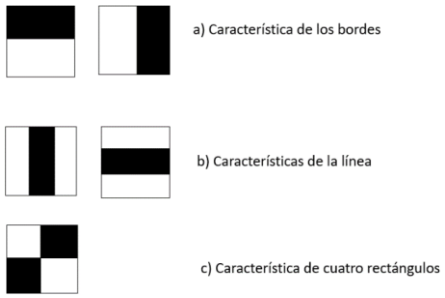


Figura 1, Características

Ahora, todos los tamaños y ubicaciones posibles de cada núcleo se usan para calcular muchas características. (¿Imagina cuánto cómputo necesita? Incluso una ventana de 24x24 da como resultado más de 160000 características). Para cada cálculo de características, debemos encontrar la suma de los píxeles debajo de los rectángulos blancos y negros. Para resolver esto, introdujeron la imagen integral. Sin importar cuán grande sea su imagen, reduce los cálculos para un píxel dado a una operación que involucra solo cuatro píxeles.

Pero entre todas estas características que calculamos, la mayoría de ellas son irrelevantes. Por ejemplo, considere la imagen de abajo. La fila superior muestra dos buenas características. La primera característica seleccionada parece centrarse en la propiedad de que la región de los ojos suele ser más oscura que la región de la nariz y las mejillas. La segunda característica seleccionada se basa en la propiedad de que los ojos son más oscuros que el puente de la nariz. Pero las mismas ventanas aplicadas a las mejillas o en cualquier otro lugar son irrelevantes. Entonces, ¿cómo seleccionamos las mejores características de las más de 160000 características? Se consigue por el algoritmo de Adaboost (Adaptive Boosting/Impulso adaptativo).[24]

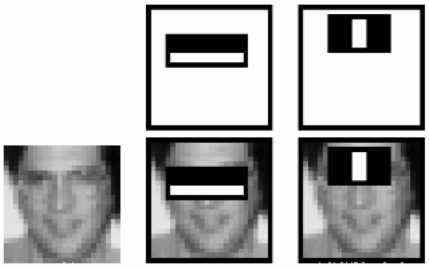


Figura 2, Identificación de características

AdaBoost un algoritmo que entrena una serie de clasificadores débiles de manera iterativa, de modo que cada nuevo clasificador por ”weak learner/Aprendiz débil” se enfoque en los datos que fueron erróneamente clasificados por su predecesor, de esta manera el algoritmo se adapta y logra obtener mejores resultados.

[24] OpenCV, computer Vision, Face Detection using Haar Cascades.

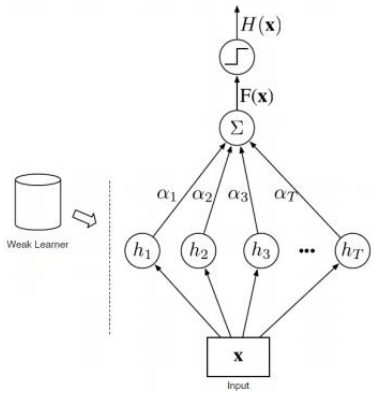


Figura 3, AdaBoost

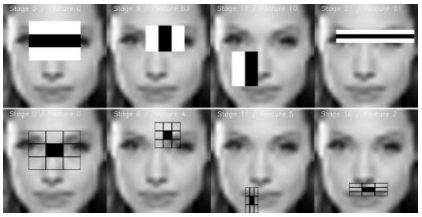


Figura 4, Visualización OpenCV, Clasificador Cascade

Cada etapa se almacena como una imagen para la validación futura de las características.

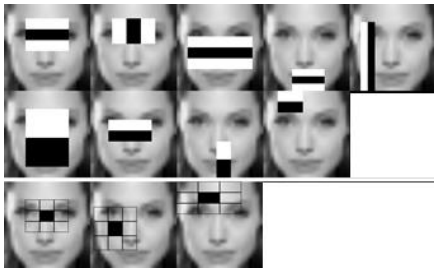


Figura 5, Clasificador Cascade

Figura 7, Detección de rostro y extracción rostro (fuente propia)

3.2 Deep Learning y TensorFlow

3.2.1 Deep Learning

Deep learning [34] es una rama de machine learning que se entrena con grandes cantidades de datos y trata con muchas unidades computacionales que trabajan en conjunto para realizar predicciones [58].

Figura 8, Inteligencia Artificial y Deep Learning (MIT- https://deeplearning.mit.edu)

3.2.2 Estructura de las Redes Neuronales Convolucionales

- Una capa convolucional, que es la que le da el nombre a la red.
- Una capa de reducción o de pooling, la cual va a reducir la cantidad de parámetros al quedarse con las características más comunes.
- Una capa clasificadora totalmente conectada, la cual nos va dar el resultado final de la red.

Los tipos de capas para redes convolucionales [18], [19], [56] son:

- Convolutional (CONV)
- Activation (ACT or RELU, where we use the same or the actual activation function)
- Pooling (POOL)

- Fully connected (FC)
- Batch normalization (BN)
- Dropout (DO)

3.2.3 La Capa de Convolución

Una convolución [18], [19], [56] es un operador matemático que transforma dos funciones f y g en una tercera función que en cierto sentido representa la magnitud en la que se superponen f y una versión trasladada e invertida de g. Una convolución es un tipo muy general de media móvil, como se puede observar si una de las funciones se toma como la función característica de un intervalo.

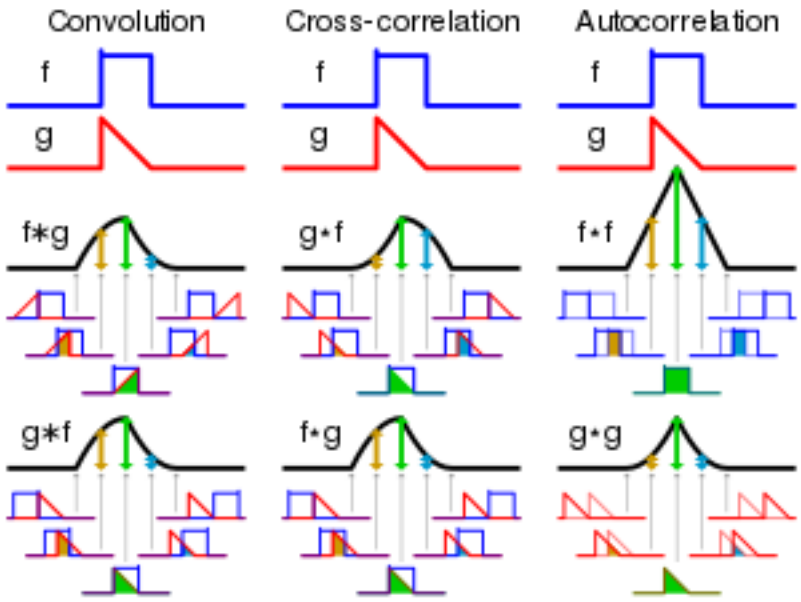


Figura 6, Convolución y correlación.

La simetría de la función f hace que g*f y f*g sean operaciones idénticas con resultados idénticos.

Convolución y correlación. La simetría de la función f hace que g*f Y f*g sean operaciones idénticas con resultados idénticos [2], [19].

Ecuación 1, Convolución y correlación

$$(f * g)(t) \doteq \int_{-\infty}^{\infty} f(\eta) g(t - \eta) d\eta$$

La siguiente figura representa gráficamente la operación de convolución, utilizando el operador (*), Output = Input * Kernel.

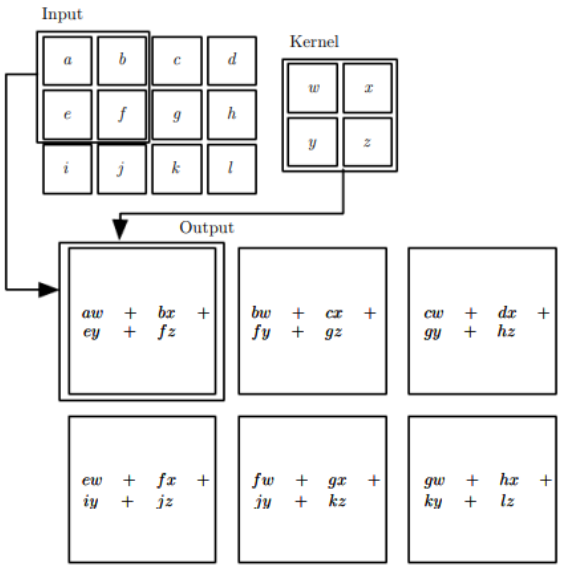


Figura 9, Simple operación convolucional

Ejemplo de la operación de convolución:

Ecuación 2, Operación discreta de convolución

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m]$$
$$= \sum_{m=-\infty}^{\infty} f[n-m]g[m].$$

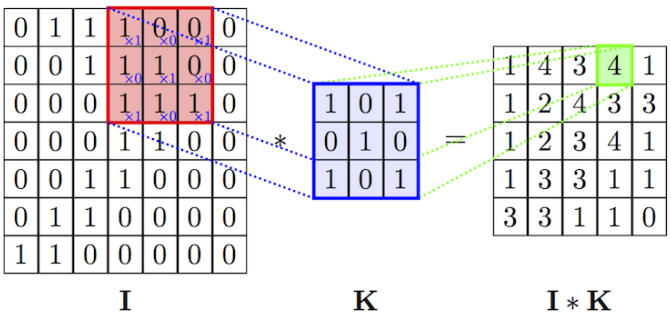


*Figura 10, la operación de convolución se indica con un asterisco (*), I*K*

La siguiente tabla muestra la operación de convolución.

Tabla 2, Ejemplo de operación de convolución

Ítem	Operación
1.	**I*K =** I[0][3]*K[0][0] + I[0][4]*K[0][1] + I[0][5]*K[0][2] + I[1][3]*K[0][0] + I[1][4]*K[1][1] + I[1][5]*K[1][2] + I[2] [3]*K[0][0] + I[2][4]*K[2][1] + I[2][5]*K[2][2]
2.	**I*K =** 1*1 + 0*0 + 0*1 + 1*0 + 1*1 + 0*0 + 1*1 + 1*0 +1*1
3.	**I*K** = 4

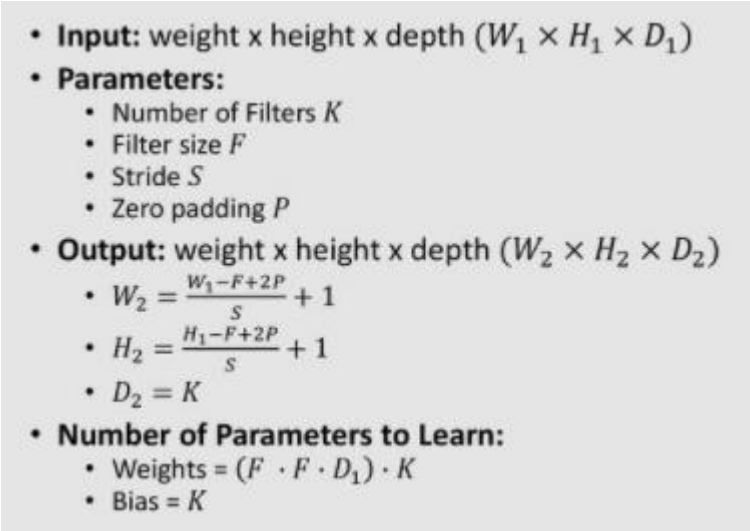


Figura 11, Operaciones en la Capa convolucional [56]

La operación de convolución [58], recibe como entrada o input la imagen y luego aplica sobre ella un filtro o kernel [56] que nos devuelve un mapa de las características de la imagen original, de esta forma logramos reducir el tamaño de los parámetros. La convolución [56], aprovecha tres ideas importantes que pueden ayudar a mejorar cualquier sistema de machine learning, ellas son:

- Interacciones dispersas, ya que al aplicar un filtro de menor tamaño sobre la entrada original podemos reducir drásticamente la cantidad de parámetros y cálculos;

- Los parámetros compartidos, que hace referencia a compartir los parámetros entre los distintos tipos de filtros, ayudando también a mejorar la eficiencia del sistema; y
- las representaciones equivariantes indican que, si las entradas cambian, las salidas van a cambiar también en forma similar.

Por otra parte, la convolución proporciona un medio para trabajar con entradas de tamaño variable, lo que puede ser también muy conveniente.

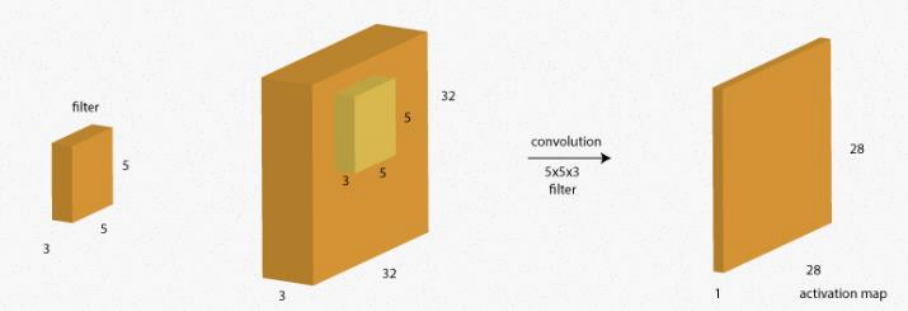


Figura 12, Filtro de convolución

3.2.4 Capa de Reducción o Pooling

La capa de reducción o pooling [56] se coloca generalmente después de la capa convolucional [56]. Su utilidad principal radica en la reducción de las dimensiones espaciales (ancho x alto) del volumen de entrada para la siguiente capa convolucional. No afecta a la dimensión de profundidad del volumen. La operación realizada por esta capa también se llama reducción de muestreo, ya que la reducción de tamaño conduce también a la pérdida de información. Sin embargo, una pérdida de este tipo puede ser beneficioso para la red por dos razones:

- La disminución en el tamaño conduce a una menor sobrecarga de cálculo para las próximas capas de la red;
- También trabaja para reducir el sobreajuste.

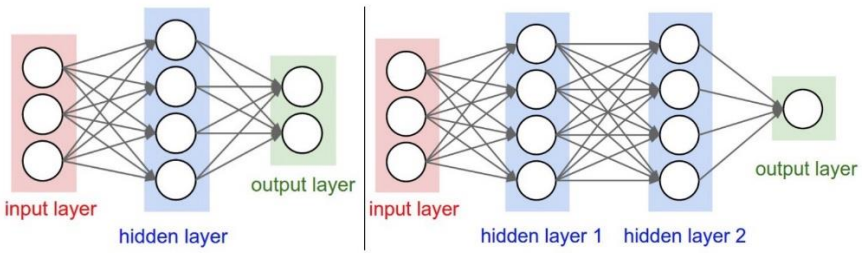


Figura 7, Las redes neuronales constan de una capa de entrada, capas ocultas y una capa de salida

(Fuente: https://bit.ly/2Hxhjaw)

La capa convolucional (o capa Conv) es el núcleo de lo que hace una red neuronal convolucional. La capa Conv consiste en un conjunto de filtros. Cada filtro se puede considerar como un cuadrado pequeño (con un ancho y una altura fijos) que se extiende a través de la profundidad total del volumen de entrada.

Durante cada pasada, el filtro 'se desplaza' en el ancho y la altura del volumen de entrada. Este proceso da como resultado un mapa de activación bidimensional que proporciona las respuestas de ese filtro en cada posición espacial.

Para evitar un ajuste excesivo, las capas de agrupación se utilizan para aplicar la reducción de muestreo no lineal en los mapas de activación. En otras palabras, agrupar capas es agresivo al descartar información, pero puede ser útil si se usa de forma adecuada. Una capa Pooling a menudo seguiría una o dos Capas Conv en la arquitectura CNN.

La operación que se suele utilizar en esta capa es **max-pooling** [58], que divide a la imagen de entrada en un conjunto de rectángulos y, respecto de cada subregión, se va quedando con el máximo valor.

Una función de agrupación reemplaza la salida de la red en una ubicación determinada con una estadística resumida de las salidas cercanas. Por ejemplo, la operación max-pooling informa el número máximo en el vecindario.

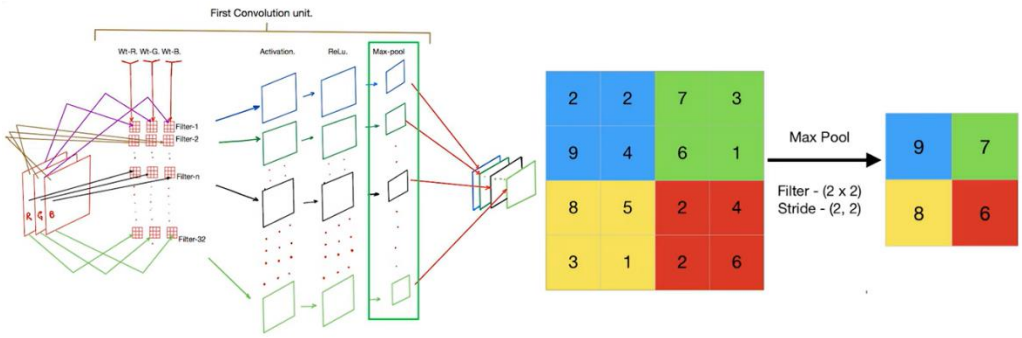


Figura 8, Corte en una sola profundidad

3.2.5 Capa clasificadora totalmente conectada

Al final de las capas convolucional y de pooling, las redes utilizan generalmente capas completamente conectados en la que cada pixel se considera como una neurona separada al igual

que en una red neuronal regular. Esta última capa clasificadora tendrá tantas neuronas como el número de clases que se debe predecir.

3.2.6 SoftMax

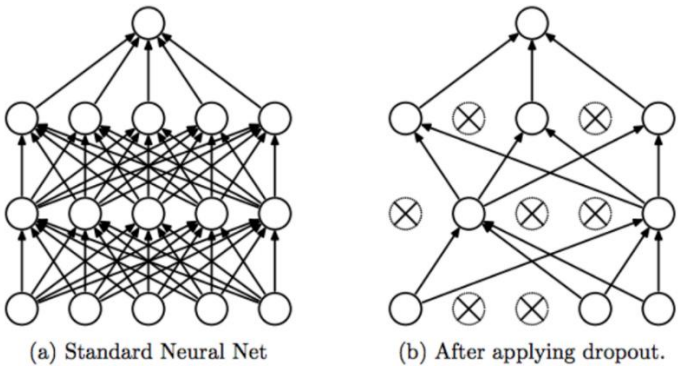
La función softmax es una función que convierte un vector de K valores reales en un vector de K valores reales que suman 1. Los valores de entrada pueden ser positivos, negativos, cero o mayores que uno, pero el softmax los transforma en valores entre 0 y 1, de modo que puedan interpretarse como probabilidades. Si una de las entradas es pequeña o negativa, el softmax la convierte en una probabilidad pequeña, y si una entrada es grande, entonces la convierte en una probabilidad grande, pero siempre permanecerá entre 0 y 1.

La función softmax se denomina a veces función softargmax o regresión logística multiclase. Esto se debe a que la softmax es una generalización de la regresión logística que puede utilizarse para la clasificación multiclase, y su fórmula es muy similar a la función sigmoidea que se utiliza para la regresión logística. La función softmax puede utilizarse en un clasificador sólo cuando las clases son mutuamente excluyentes.

Muchas redes neuronales multicapa terminan en una penúltima capa que emite puntuaciones de valor real que no están convenientemente escaladas y con las que puede ser difícil trabajar. En este caso, el softmax es muy útil porque convierte las puntuaciones en una distribución de probabilidad normalizada, que puede mostrarse a un usuario o utilizarse como entrada para otros sistemas. Por eso es habitual añadir una función softmax como capa final de la red neuronal [66].

3.2.7 Dropout

El abandono es en realidad una forma de regularización que pretende ayudar a prevenir el sobreajuste aumentando la precisión de las pruebas, quizás a expensas de la precisión del entrenamiento. Para cada mini lote de nuestro conjunto de entrenamiento, las capas de abandono, con probabilidad p, desconectan aleatoriamente las entradas de la capa anterior a la siguiente capa de la arquitectura de la red [9].



3.2.8 Arquitectura CNN

La siguiente imagen muestra la representación de la arquitectura de una red neuronal convolucional [9].

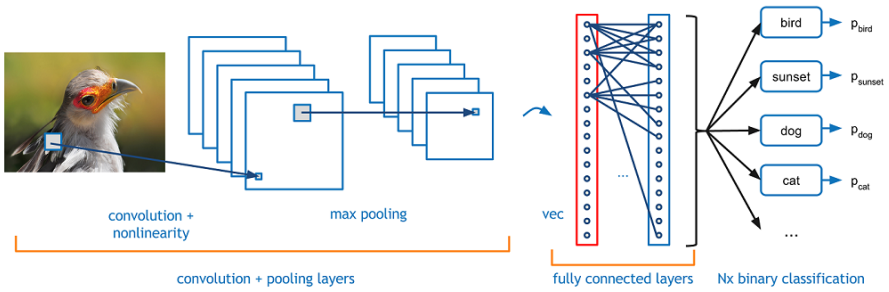


Figura 9, Un ejemplo de CNN Layer Architecture for Image Classification

(Fuente: https://bit.ly/2vwlegO)

La representación base de las imágenes están dadas por pixeles, que representan la combinación de los tres colores básicos RGB (red, green, blue).

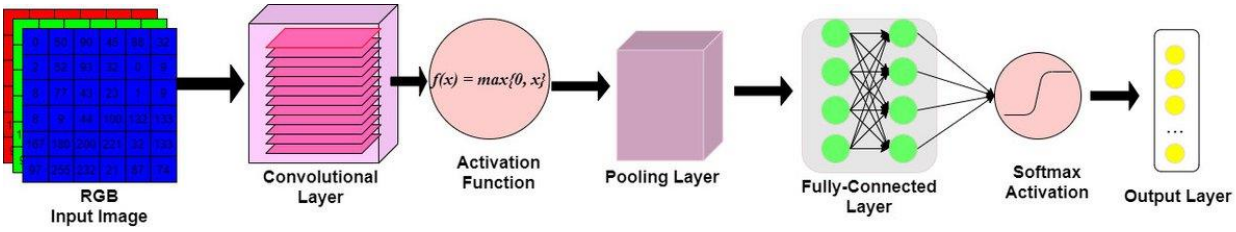


Figura 13, Arquitectura de una red neuronal convolucional

3.2.9 Funciones de Activación

La salida de las capas y de la red neuronal se procesa utilizando una función de activación, que es un nodo que se agrega a las capas de ocultación y a la capa de salida [9].

A menudo encontrará que la función de activación de ReLu se usa en capas ocultas, mientras que la capa final generalmente consiste en una función de activación de SoftMax. La idea es que, al apilar capas de funciones lineales y no lineales, podemos detectar una amplia gama de patrones y predecir con precisión una etiqueta para una imagen determinada [17].

SoftMax a menudo se encuentra en la capa final que actúa básicamente como un normalizador y produce un vector discreto de distribución de probabilidad, que es genial para nosotros, ya que la salida de la CNN [17] que queremos es una probabilidad de que una imagen se corresponda con una clase en particular.

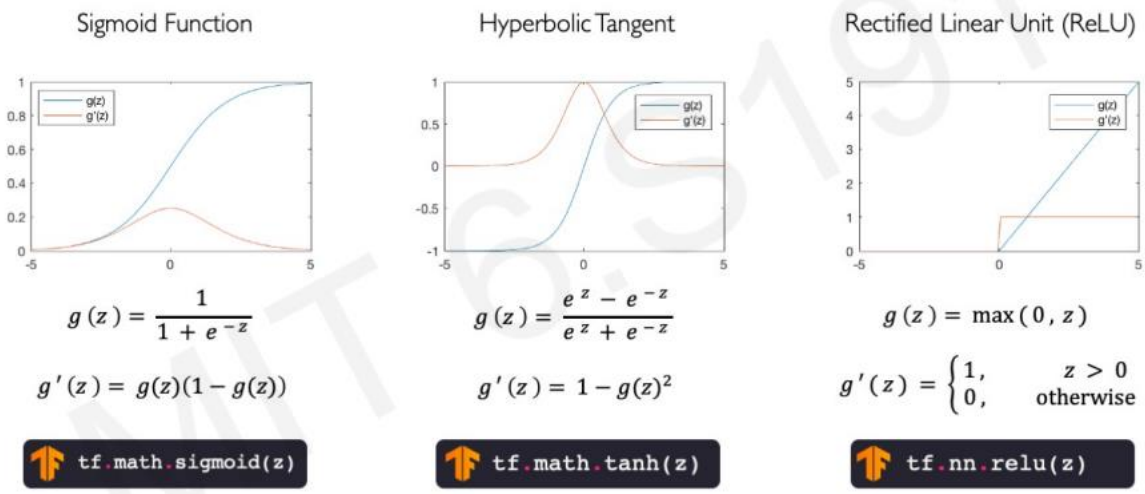


Figura 14, Las funciones de activación en Tensorflow (fuente 6S191_MIT_DeepLearning_L1)

3.2.10 Tensorflow y Tensores

TensorFlow [10], como su nombre lo indica, es un marco para definir y ejecutar cálculos que involucran tensores. Un tensor es una generalización de vectores y matrices a dimensiones potencialmente más altas. Internamente, TensorFlow representa los tensores como matrices n-dimensionales de tipos de datos base.

TensorFlow es una interfaz para expresar algoritmos de aprendizaje automático y una implementación para ejecutar dichos algoritmos.

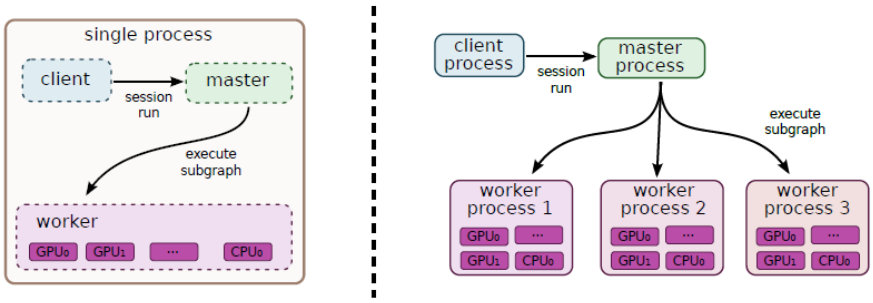


Figura 15, Estructura de una máquina y de un sistema distribuido [10]

Al escribir un programa TensorFlow[11], el objeto principal que manipula y transmite es el tf.Tensor. Un objeto tf.Tensor representa un cálculo parcialmente definido que eventualmente producirá un valor. Los programas TensorFlow funcionan construyendo primero un gráfico de los objetos de Tf.Tensor, detallando cómo se calcula cada tensor basándose en los otros tensores disponibles y luego ejecutando partes de este gráfico para lograr los resultados deseados.

Los tipos de datos (float32, int32 o cadena).
Cada elemento en el Tensor tiene el mismo tipo de datos, y el tipo de datos siempre se conoce. La forma (es decir, el número de dimensiones que tiene y el tamaño de cada dimensión) puede ser conocida solo parcialmente. La mayoría de las operaciones producen tensores de formas completamente conocidas si las formas de sus entradas también son completamente conocidas, pero en algunos casos solo es posible encontrar la forma de un tensor en el momento de la ejecución del gráfico.

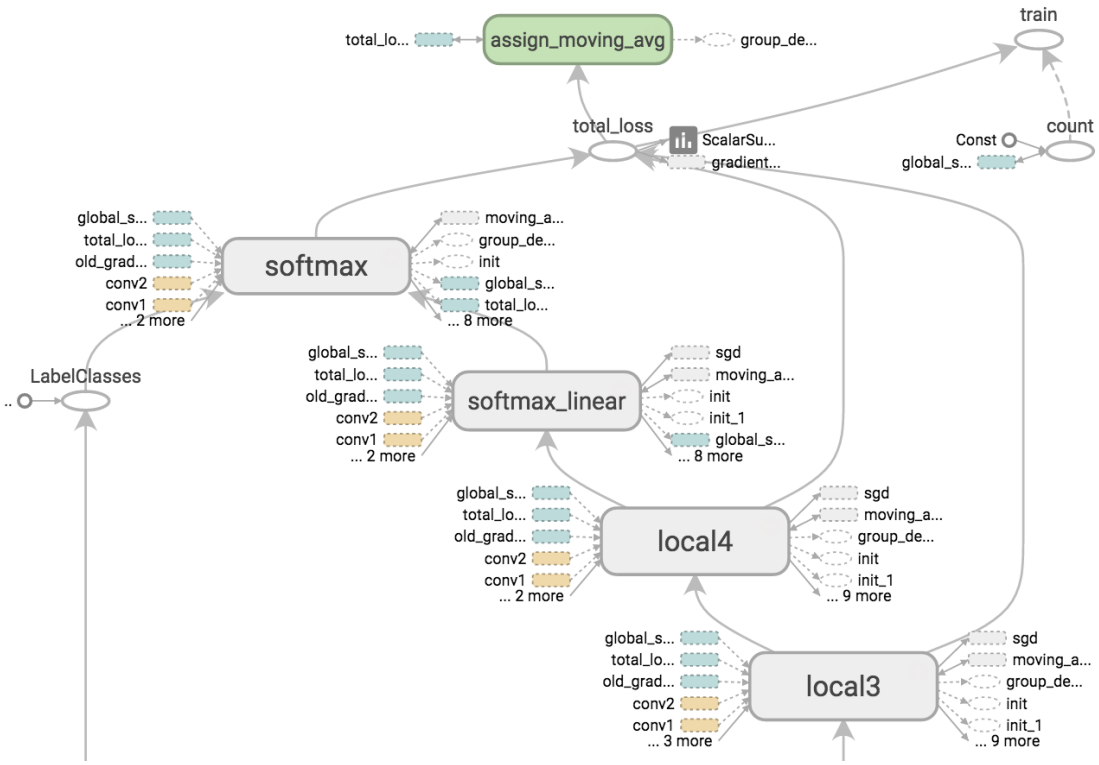


Figura 16, Visualización gráfica en TensorBoard de un modelo de red neuronal convolucional [10],

Algunos tipos de tensores son especiales, Los principales son [11]:

- tf.Variable
- tf.constant
- tf.placeholder
- tf.SparseTensor

Shape, La forma de un tensor es el número de elementos en cada dimensión. TensorFlow infiere automáticamente formas durante la construcción del gráfico. Estas formas inferidas pueden tener rango conocido o desconocido. Si se conoce el rango, los tamaños de cada dimensión pueden ser conocidos o desconocidos.

La documentación de TensorFlow utiliza tres convenciones de notación para describir la dimensionalidad del tensor: rango, forma y número de dimensión. La siguiente tabla muestra cómo se relacionan entre sí:

Rank	Shape	Dimension number	Example
0	[]	0-D	A 0-D tensor. A scalar.
1	[D0]	1-D	A 1-D tensor with shape [5].
2	[D0, D1]	2-D	A 2-D tensor with shape [3, 4].
3	[D0, D1, D2]	3-D	A 3-D tensor with shape [1, 4, 3].
n	[D0, D1, ... Dn-1]	n-D	A tensor with shape [D0, D1, ... Dn-1].

Figura 10, Relación de tensores y shape[25]

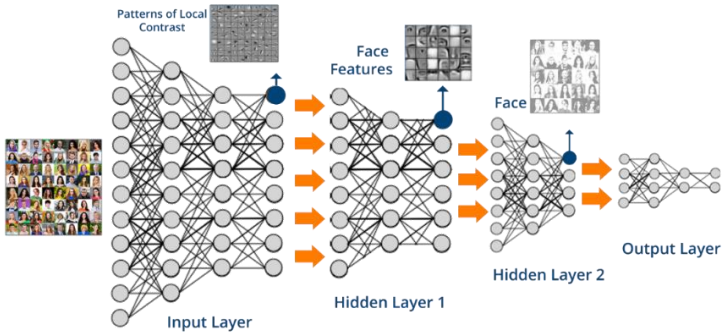


Figura 11, Modelo de CNN

Las aplicaciones de Keras son modelos de aprendizaje profundo que están disponibles junto con pesos previamente entrenados. Estos modelos se pueden utilizar para la predicción, la extracción de características y el ajuste fino.

Tabla 3, Modelos disponibles KERAS, número de parámetros, precisión y tiempo (fuente keras.io)

Modelo	Tamaño (MB)	Precisión Top-1	Precisión entre las 5 mejores	Parámetros	Profundidad	Tiempo (ms) por paso de inferencia (CPU)	Tiempo (ms) por paso de inferencia (GPU)
Xception	88	0,790	0,945	22,910,480	126	109,42	8.06
VGG16	528	0,713	0.901	138,357,544	23	69,50	4.16
VGG19	549	0,713	0.900	143,667,240	26	84,75	4.38
ResNet50	98	0,749	0,921	25,636,712	-	58,20	4.55

[25] https://www.tensorflow.org/guide/tensors

La precisión top 1 y top 5 se refiere al rendimiento del modelo en el conjunto de datos de validación de ImageNet.

La profundidad se refiere a la profundidad topológica de la red. Esto incluye capas de activación, capas de normalización de lotes, a continuación, mostramos cómo se adecuan a la arquitectura de TensorFlow.

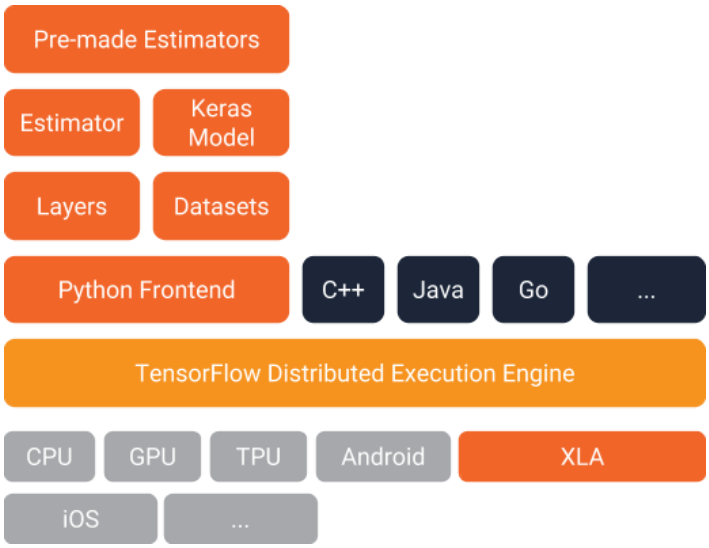


Figura 12, Modelo de TensorFlow & Keras

TensorFlow es una herramienta de machine learning. Popularizada por su eficiencia con redes neuronales de aprendizaje profundo pero que permite la ejecución de procesos distribuidos que no tengan nada que ver con redes neuronales. los Estimators representan un modelo completo y ofrecen los mecanismos para entrenar los modelos, evaluar su error y generar predicciónes. Son la capa de más alto nivel.

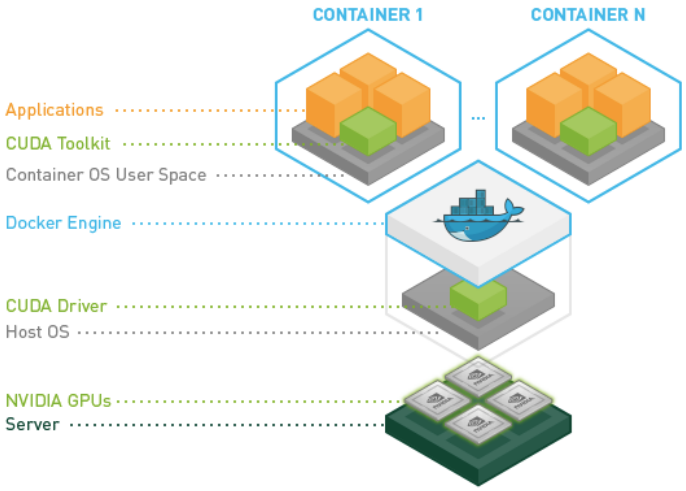


Figura 13, Modelo de TensorFlow usando Contenedores[10]

Nvidia-Docker es básicamente una envoltura alrededor de la CLI de la ventana acoplable que aprovisiona de forma transparente un contenedor con las dependencias necesarias para ejecutar código en la GPU.

3.2.11 Big Data

Big data, o los grandes volúmenes de datos, es un término que describe cualquier cantidad voluminosa de datos estructurados, semiestructurados y no estructurados, que tienen el potencial de ser extraídos para obtener información. A este respecto Gartner, la consultora de tecnologías de información, define que los datos no estructurados son aquellos que no pertenecen a un modelo no definido o específico. Incluso estima que el 80 por ciento de los datos son no estructurados, como el cuerpo de un correo. En pocas palabras, los datos no estructurados son aquellos que no caben en una base de datos. Por su parte, los datos estructurados son aquellos que pueden ser ingresados en una base, como: fecha, hora, nombres, cuentas de correo, números, etc. En pocas palabras, hay un sinfín de fuentes de datos que producen información nueva cada minuto.

Actualmente, la adopción de soluciones de big data permite que las compañías ocupen datos para su beneficio. Para lograrlo, primero hay que considerar las 5 Vs de big data: el Volumen extremo de datos, la gran Variedad de tipos de datos y la Velocidad a la que se deben procesar dichos datos, la Veracidad de los datos, y lo más importante el Valor de los datos.

Volumen de datos. La organización debe poder extraer un conjunto de datos de entrenamiento que pueda analizarse rápidamente utilizando los diferentes algoritmos candidatos, pero también uno

que refleje adecuadamente el conjunto completo de datos. Su organización debe contar con suficientes datos para representar correctamente las incidencias reales, a la vez que considera todos los escenarios posibles para los resultados que está buscando.

Variedad de datos. Conforme los datos no estructurados se vuelven esenciales para los procesos de negocio, se convierten en fuentes necesarias para los modelos predictivos. Esto significa que una empresa debe tener un conjunto de procesos sólidos para escanear, analizar y contextualizar datos no estructurados y transformarlos en conjuntos de datos que alimenten los algoritmos de analítica.

Velocidad de los datos. Cada vez más rápido se reciben grandes cantidades de datos y a menudo no se puede predecir cuándo podrían cambiar, lo que forzaría una necesidad casi continúa de creación y preparación de perfiles de datos. La necesidad de grandes velocidades de datos requiere una potente infraestructura de cómputo, capaz de procesar rápidamente grandes volúmenes y variedades de datos sin sobrecargar los servidores. Esto puede requerir cientos o miles de servidores para distribuir el trabajo y operar de manera colaborativa. Pero la computación en la nube ofrece alternativas para el almacenamiento y análisis de datos, de forma que las empresas puedan implementar proyectos de big data sin impactar su propia infraestructura.

Figura 14, Modelo de BigData

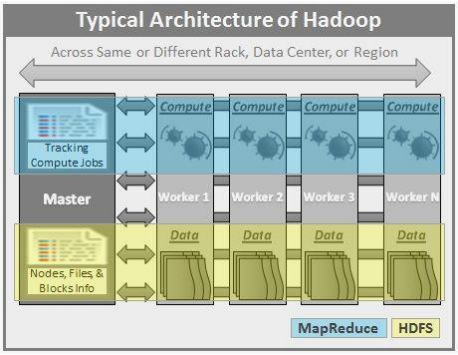


Figura 15, Arquitectura Apache Hadoop Cluster

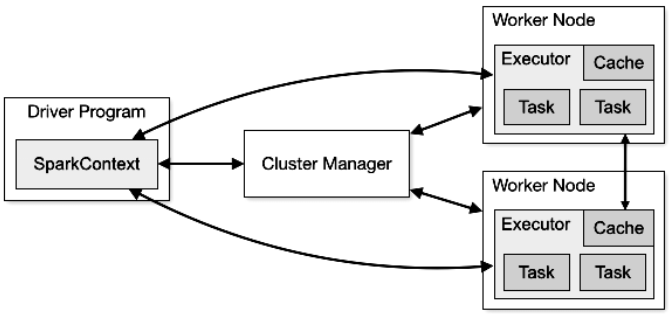


Figura 16, Apachde Spark Cluster

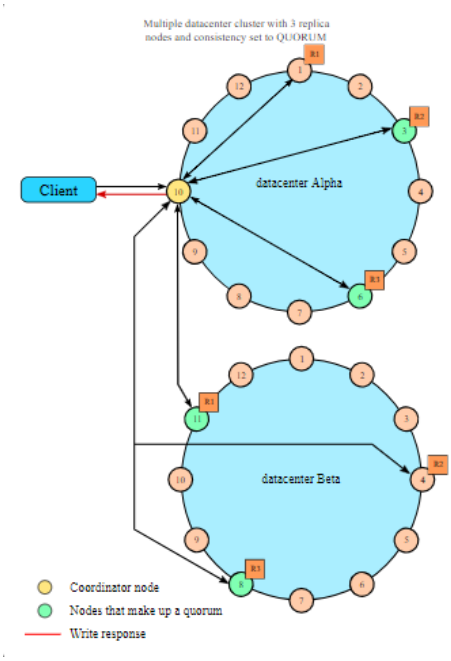


Figura 17, Arquitectura Apache Cassandra mutiple datacenter [12]

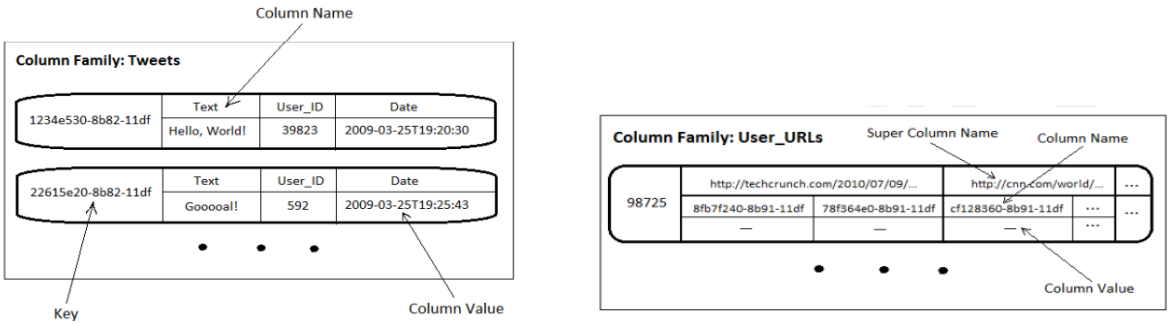
Cassandra [12], se organiza en familias de columnas.

A continuación, los componentes de una base de datos Cassandra y la terminología utilizada:

- **Espacio de claves**. Es el término utilizado para referirnos a una base de datos. Un espacio de claves es una colección ordenada de familia de columnas.
- **Familia de columna**. Es una colección ordenada de columnas. Podríamos asociarlo al concepto clásico de tabla utilizado en las bases de datos relacionales.
- **Clave de fila**. Identificador único de un dato almacenado dentro de una familia de columnas. Asociable a la primary key de las bases de datos relacionales

- **Columnas**. Es la unidad básica del modelo de datos formado por una combinación de clave, valo y timestamp.
- **SuperColumna**. Es una columna cuyos valores son una o más columnas,denominadas subcolumnas, que se encuentran ordenadas.

Tabla 4, Vista del modelo de datos Cassandra [10]



3.2.12 Matriz de Confusión

En el campo del aprendizaje automático y específicamente el problema de la clasificación estadística, una matriz de confusión, también conocida como matriz de error, es un diseño de tabla específico que permite la visualización del desempeño de un algoritmo, típicamente uno de aprendizaje supervisado (en aprendizaje no supervisado por lo general se denomina matriz de emparejamiento). Cada fila de la matriz representa las instancias en una clase real, mientras que cada columna representa las instancias en una clase predicha, o viceversa; ambas variantes se encuentran en la literatura. El nombre proviene del hecho de que facilita ver si el sistema está confundiendo dos clases (es decir, comúnmente etiquetado incorrectamente una como otra).

Terminología de una matriz de confusión:

Terminología	Descripción
condición positiva (P)	el número de casos positivos reales en los datos
condición negativa (N)	el número de casos negativos reales en los datos
verdadero positivo (TP)	Equivalente con aceptación
verdadero negativo (TN)	Equivalente con correcto rechazo

falso positivo (FP)	Equivalente con falsa alarma , error tipo I o subestimación
falso negativo (FN)	Equivalente con falla, error o sobreestimación

		Predicción	
		Positivos	**Negativos**
Observación	**Positivos**	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Figura 17, Matriz de confusión

la matriz de confusión sigue las siguientes reglas de la figura 18, evaluando el desempeño de las modelos entrenadas, considerando los falsos positivos, falsos negativos, verdaderos positivos y verdaderos negativos [65].

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})}$$

$$\text{Recall} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})}$$

$$\text{F1-score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{(\text{True Positive} + \text{False Positive} + \text{True Negative} + \text{False Negative})}$$

Figura 18, Evaluación de desempeño de modelos entrenados[65]

3.2.13 Interoperabilidad

El Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) define interoperabilidad como la habilidad de dos o más sistemas o componentes para intercambiar información y utilizar la información intercambiada [5].

El ámbito de la administración electrónica ha dotado a la interoperabilidad [5] de gran relevancia y ha impulsado estudios científicos que actualmente destacan otras dimensiones por encima de la dimensión técnica de la interoperabilidad [5].

Es precisamente en este contexto donde se impone la interoperabilidad actualmente como uno de los elementos clave para la administración electrónica, reflejado en España y México a través del Esquema Nacional de Interoperabilidad.

Además de hablarse de la gobernanza de la interoperabilidad, se reconoce a la interoperabilidad actualmente, al menos, tres dimensiones bien diferenciadas:

Dimensión técnica, donde se especifica todos los elementos técnicos la implementar, desde la parte de infraestructura, componentes de software, componentes de seguridad y de integración continua.

Dimensión semántica, establece todas las normativas establecidas para estandarizar los mecanismos de comunicación entre las instituciones públicas y el ciudadano.

Dimensión organizacional, las jerarquías y roles organizacionales de las entidades públicas.

Factores de Influencia, El contexto del gobierno electrónico es complejo porque tiene que lidiar con aspectos legales, y políticos, y Cuestiones socioculturales. Por ejemplo, la naturaleza independiente de las agencias gubernamentales y los conflictos El liderazgo, la política y las prioridades financieras dificultan la consolidación y la cooperación. en la búsqueda de mejores servicios para los ciudadanos. En nuestro modelo de interoperabilidad propuesto, representamos este impacto adicional de política, política (legal), y las cuestiones socioculturales como factores influyentes. Estos factores son capturados de esta manera y no Como nivel adicional por estas razones:

- El impacto de cada uno es diferente, dependiendo de la situación.
- Se aplican a casi cualquier sistema de gobierno electrónico, independientemente de la interoperabilidad prevista nivel.

Por ejemplo, incluso un sitio web simple que comparte información con el público debe considerar legal cuestiones tales como qué información se considera pública, cuestiones políticas como la coherencia con políticas y objetivos actuales, y cuestiones socioculturales como la accesibilidad para personas que no están en inglés altavoces Desde la perspectiva del desarrollador de sistemas, estos factores se convierten en una dimensión adicional y crítica de El problema de la interoperabilidad [5] y cualquier esfuerzo de gobierno electrónico deben abordarlos. Alcanzando cualquier interoperabilidad El nivel requiere soluciones para cumplir con los factores legales,

políticos y socioculturales. Como se muestra en el modelo de interoperabilidad propuesto en la Figura 18.

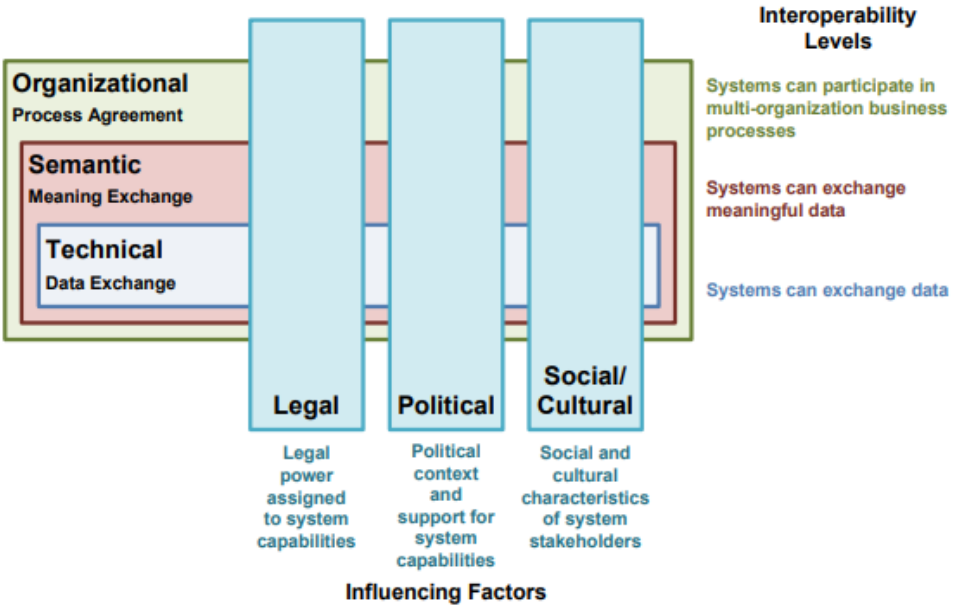


Figura 19, Modelo de interoperabilidad de gobierno electrónico [5]

CAPÍTULO 4

DESARROLLO

En este capítulo se explica la propuesta de la solución de identificación biométrica única, desarrollando las diferentes vistas de arquitectura, así como la recolección de las imágenes, su extracción, entrenamiento, y su validación utilizando la matriz de confusión. Finalmente se describe los aspectos de infraestructura y software y librerías como OpenCV, tensorflow, keras utilizados para su implementación.

4.1 Diseño de la Propuesta

Para la propuesta se inicia un mecanismo de registro de la foto del rostro del ciudadano habilitando la video cámara, y así un ciudadano que vive en la ciudad de Arequipa y requiere realizar un trámite en Cajamarca, solo requerirá una video cámara y desde su domicilio y por internet podrá acceder a realizar su trámite en línea y de forma remota con un funcionario de gobierno con previa identificación biométrica, es importante considerar la seguridad, un servidor Streaming así como la habilitación de la 5G[26].

La interoperabilidad [5] es un tema importante para las instituciones de gobierno, actualmente los ciudadanos realizan sus trámites de forma presencial, generando pérdida de tiempo y por ende de dinero.

Identificar a un ciudadano a partir de sus características biométricas(rostro) para que pueda autenticarse con la plataforma de interoperabilidad del estado y realizar sus trámites desde su domicilio y en línea, es posible gracias a las diferentes interfaces biométricas que se integran. Avances como ALICEN(Francia), una aplicación móvil con autenticación biométrica tiene la capacidad de poder realizar más de 500 trámites en línea, son cosas que están sucediendo en el

[26] Europa dando la visión a la 5G, https://digital-strategy.ec.europa.eu/en/policies/5g, huawei 5G https://www.huawei.com/minisite/5g-ultra-lean-site-2019/pdf_v1.0/5G-Ultra-Lean-Site-White-Paper_en.pdf

presente y nos dan indicios de lo que ocurrirá en el futuro, ¿Cómo será la autenticación para realizar trámites dentro de 5 años?, talvez será por este mecanismo biométrico.

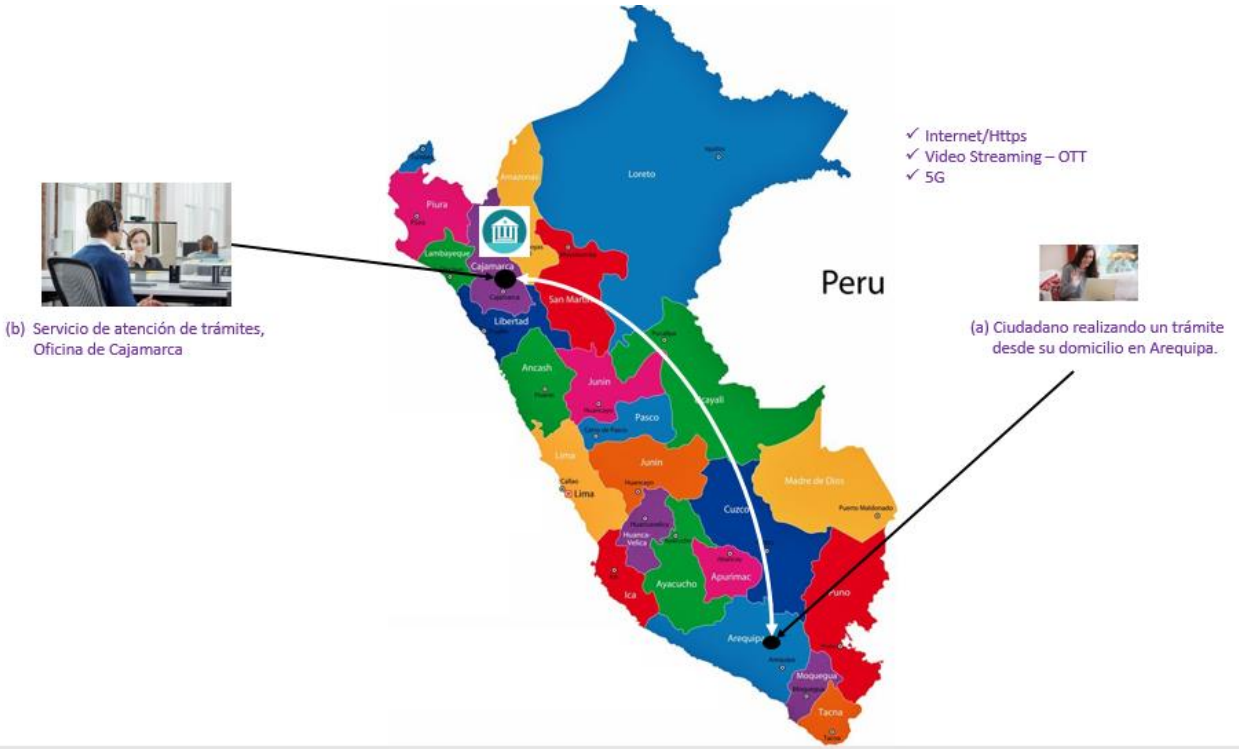


Figura 20, Diseño de la propuesta para acceder a realizar un trámite de forma remota

(fuente propia y https://www.orangesmile.com/guia-turistica/peru/mapa-pais-regiones.htm)

4.2 Recolección y Sistematización de Recolección de Conjunto de Imágenes

Se desarrolló una modulo web para poder acceder a la video cámara, capturando imágenes por demanda, y almacenándolas como primer evento, y posteriormente utilizando técnicas de clasificadores en cascada librería implementada en OpenCV (haarcascade_frontalface).

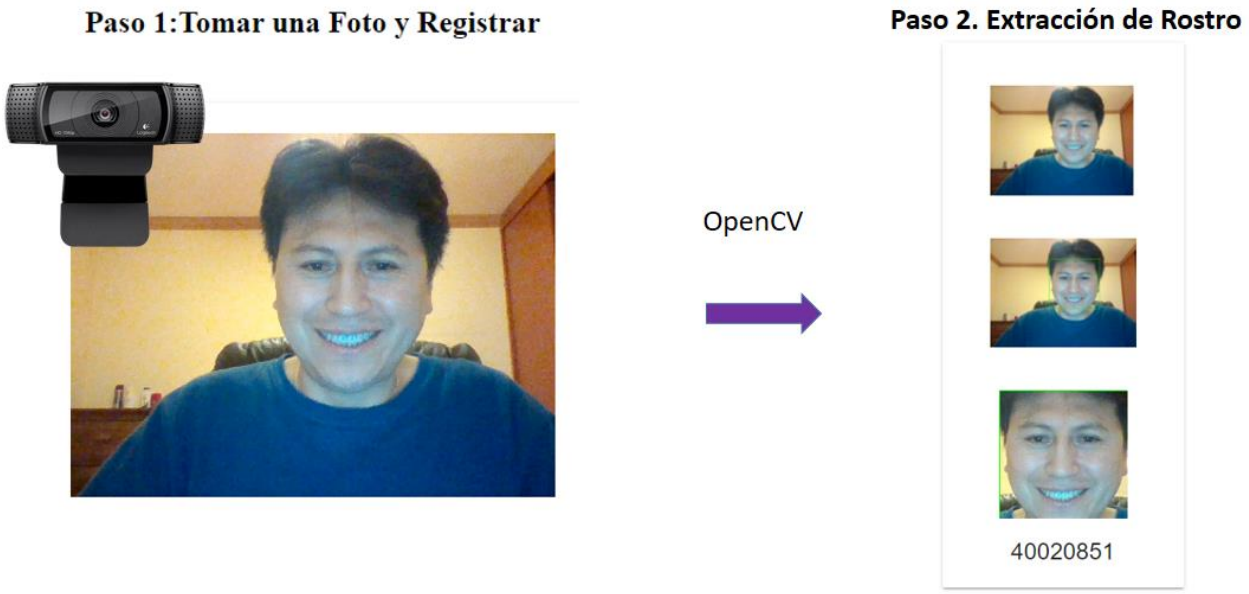


Figura 21, registro y extracción de la foto del rostro del ciudadano (fuente propia)

La siguiente tabla muestra los datos utilizados en la recolección de las imágenes(rostros).

Tabla 5, Conjunto de datos usados en la propuesta (fuente propia)

Clases	**Rostro**	**Resolución en pixeles**	**Cantidad de imágenes por clase**
1. Alfredo		224 x 224	300
2. Héctor		224 x 224	300
3. Jaime		224 x 224	300
4. Jenny		224 x 224	300
5. Jonathan		224 x 224	300

6. Martin		224 x 224	300
7. Omar		224 x 224	300
8. Roberto		224 x 224	300
9. Román		224 x 224	300

Para la extracción se consideraron 300 imágenes con por cada individuo y con una resolución de 224x224 pixeles, la población total fue de 9 individuos.

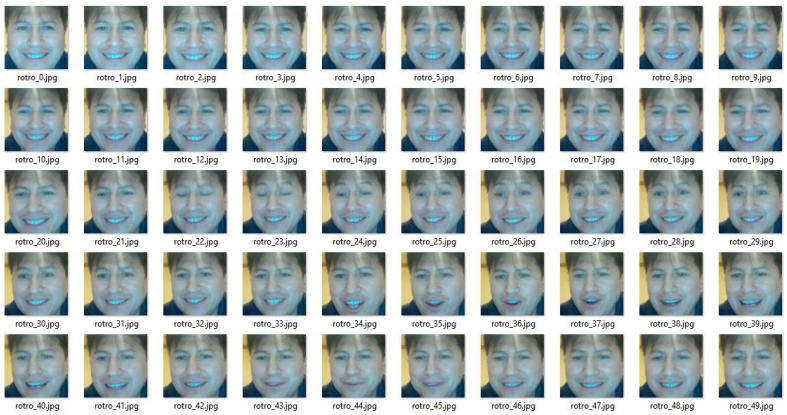


Figura 22, Ejemplo de una muestra de 50 de las 300 imágenes recolectadas (224x224 resolución por imagen - fuente propia)

4.2.1 Registro de identidad Biométrica Única del ciudadano

Tabla 6, Registro de identidad Biométrica Única del Ciudadano

Nombre	Registro de identidad Biométrica Única del ciudadano
Precondiciones	El ciudadano debe contar con su documento de identidad. El ciudadano requiere su identidad biométrica.

	El ciudadano acude a una institución de registro de identidad biométrica Única. El ciudadano lleva un USB.
Flujo Básico	1. El funcionario solicita documento de identidad. 2. El funcionario crea un nuevo registro biométrico, registra 5 imágenes del ciudadano (rasgos biométricos - rostros) y su identificador de su identidad. 3. El sistema registra en el bigdata las imágenes obtenidas del registro. 4. El sistema inicia el entrenamiento de las imágenes. 5. El funcionario realiza una validación de su identidad biométrica. 6. Si la validación es correcta. ✓ El Sistema genera certificados de autenticación. 7. El funcionario copia en el USB del ciudadano las credenciales de autenticación. 8. Si la validación es negativa ✓ El funcionario retorna en la aplicación a tomar nuevas fotos. 9. Termina el registro de identidad biométrica Única.
Flujo Alternativo	

4.2.1.1 Narrativa

El ciudadano acude a la institución de registro de identidad biométrica única, donde presentara su documento de identidad (DNI, Pasaporte, Boleta Militar) y un USB (Universal Serial Bus), el funcionario le toma las respectivas fotos y registra su identidad biométrica, posteriormente realiza la validación y luego le devuelve al ciudadano su USB con las credenciales de autenticación y termina el registro.

4.2.1.2 Proceso: Registro de identidad Biométrica Única del ciudadano

Las capacidades de detección del rostro, extracción de las características biométricas y su respectiva identificación [14], es posible gracias a diferentes componentes como son las librerías OpenCV que nos ayudara en primera instancia en capturar esas características biométricas más relevantes de un rostro.

Una vez extraído las características biométricas, esta se convertirá en la entrenada de la red neuronal convolucional implementada con las librerías de Tensorflow y Keras.

Finalmente, un ciudadano inicia sesión en un portal de una institución y esta aplicación para iniciar sesión solicitara su verificación biométrica(rostro) desde una webcam, y si es válida sus credenciales biométricas el podrá realizar el trámite en línea y podrá ver sus datos históricos de otros trámites.

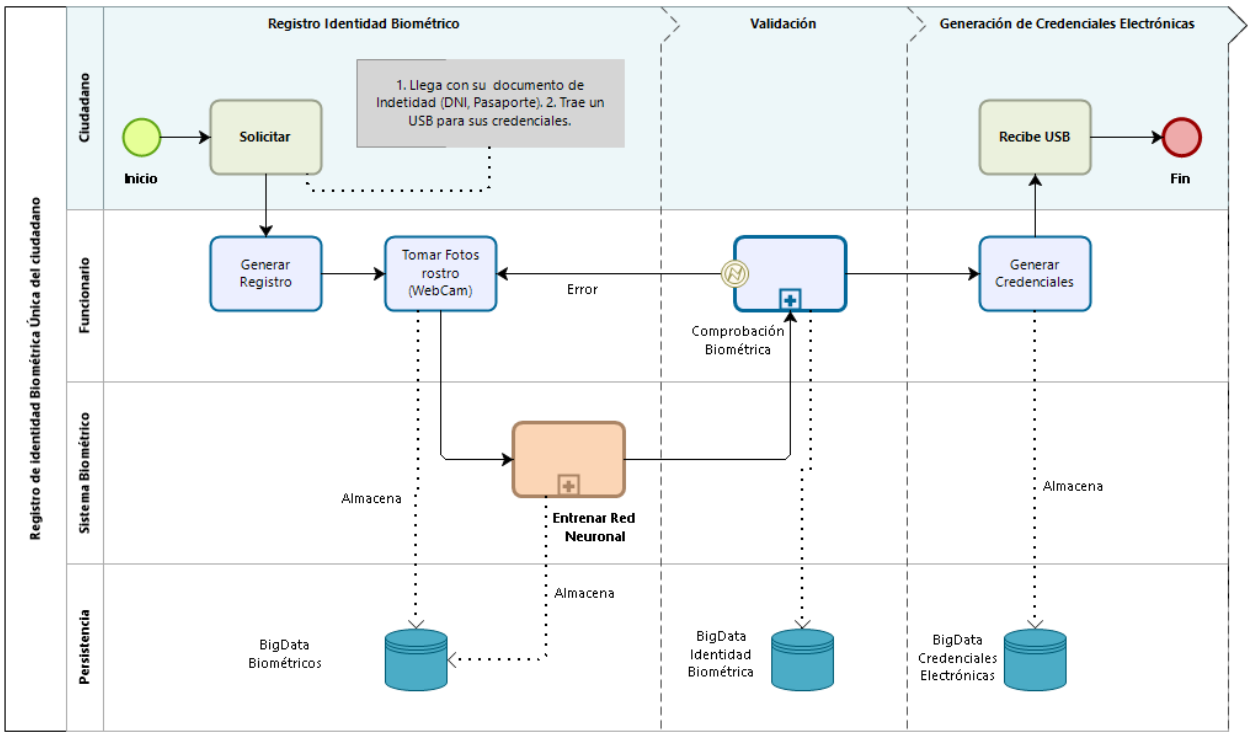


Figura 23, Registro de identidad Biométrica Única del ciudadano

4.3 Entrenamiento de Redes Neuronales Convolucionales y su Evaluación

Para realizar el entrenamiento se considera una base de datos de 2700 imágenes, la siguiente figura nos muestra el modelo, generando dos bases de datos de imágenes de forma aleatoria el primero 80% que servirá para el proceso de entrenamiento y otra base de datos de imágenes del 20% para realizar las pruebas de validación y ajustar la precisión.

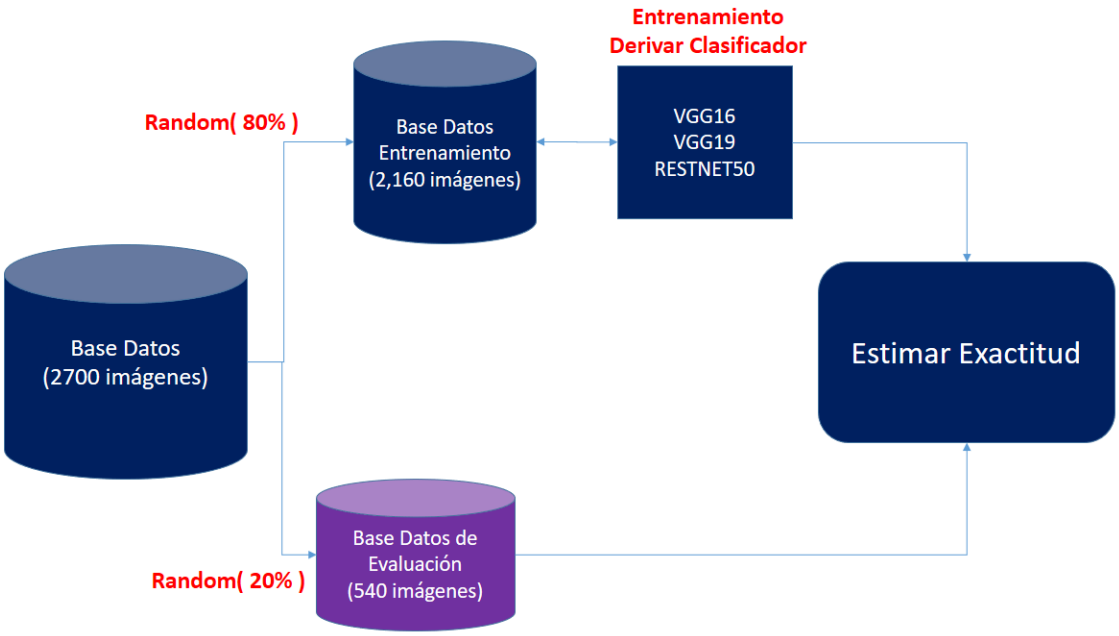


Figura 24, Modelo de entrenamiento de la CNN (fuente propia)

Para el caso del entrenamiento, se considera tres diferentes arquitecturas de redes convolucionales y son VGG16, VGG19 y el RESTNET50.

En la siguiente figura se muestra como las diferentes clases de entrada (imágenes de rostros), son entrenados utilizando los diferentes modelos como el VGG16, VGG19, RESTNET50.

Figura 25, Modelo de entrenamiento de las redes neuronales convolucionales (fuente propia)

La siguiente tabla muestra la configuración para el entrenamiento de las diferentes redes neuronales convolucionales reutilizando la arquitectura de los modelos VGG16, VGG19 y el RESTNET50.

Cada imagen tiene una resolución de 224x224 píxeles, y están en formado JPG.

Tabla 7, Configuración de las redes neuronales para su entrenamiento

Arquitectura CNN	**Total de Clases**	**Total de imágenes x Clase**	**Total de Imágenes utilizadas para entrenamiento (80%)**	**Total de Imágenes utilizadas para pruebas (20%)**	**Total de Imágenes utilizadas para evaluación (20%)**	**Total Épocas**
VGG16	9	300	240	60	60	10
VGG19	9	300	240	60	60	10
RESTNET50	9	300	240	60	60	10
VGG16 + 3	9	300	240	60	60	5

El proceso de validación del modelo está relacionado con la autenticación biométrica del ciudadano.

4.3.1 Autenticación e identificación Biométrica

Tabla 8, Autenticación de Identidad Biométrica Única del Ciudadano

Nombre	Autenticación biométrica del Ciudadano
Precondiciones	El ciudadano debe contar con su documento de identidad. El ciudadano requiere realizar un trámite en line. El ciudadano tiene su USB con credenciales electrónicas.
Flujo Básico	1. El ciudadano ingresa a gob.pe identificando su trámite. 2. El portal del trámite solicita sus credenciales electrónicas. 3. El portal realiza la validación de credenciales. 4. Si la validación es correcta. ✓ El portal solicita la validación Electrónica, iniciando cámara (WebCam) ✓ Si la validación es correcta: ✓ El ciudadano puede iniciar su trámite y ver su información histórica. ✓ Caso contrario no podrá ingresar al portal para realizar su trámite. 5. Si la validación de las credenciales electrónicas es incorrecta, no podrá continuar con el proceso de validación biométrica. 6. Termina la autenticación biométrica Única.
Flujo Alternativo	El portal identifica que el ciudadano pueda autenticarse en un único punto.

4.3.1.1 Narrativa

El ciudadano requiere realizar un trámite desde el portal de gob.pe en línea, el ciudadano tiene físicamente el USB con las credenciales electrónicas de autenticación, el sistema le pide estas credenciales y el ciudadano las carga al sistema y se valida dichas credenciales (vigencia, número de serie), si son correctas el sistema le solicita que habilite la cámara de la WebCam, una vez habilitada esta le solicitara validar su identidad biométrica(rostro).

Una vez confirmada la validación biométrica el ciudadano ingresa al portal donde podrá realizar su trámite y ver la información historia de trámites de su perfil.

4.3.1.2 Proceso: Autenticación e Identificación Biométrica

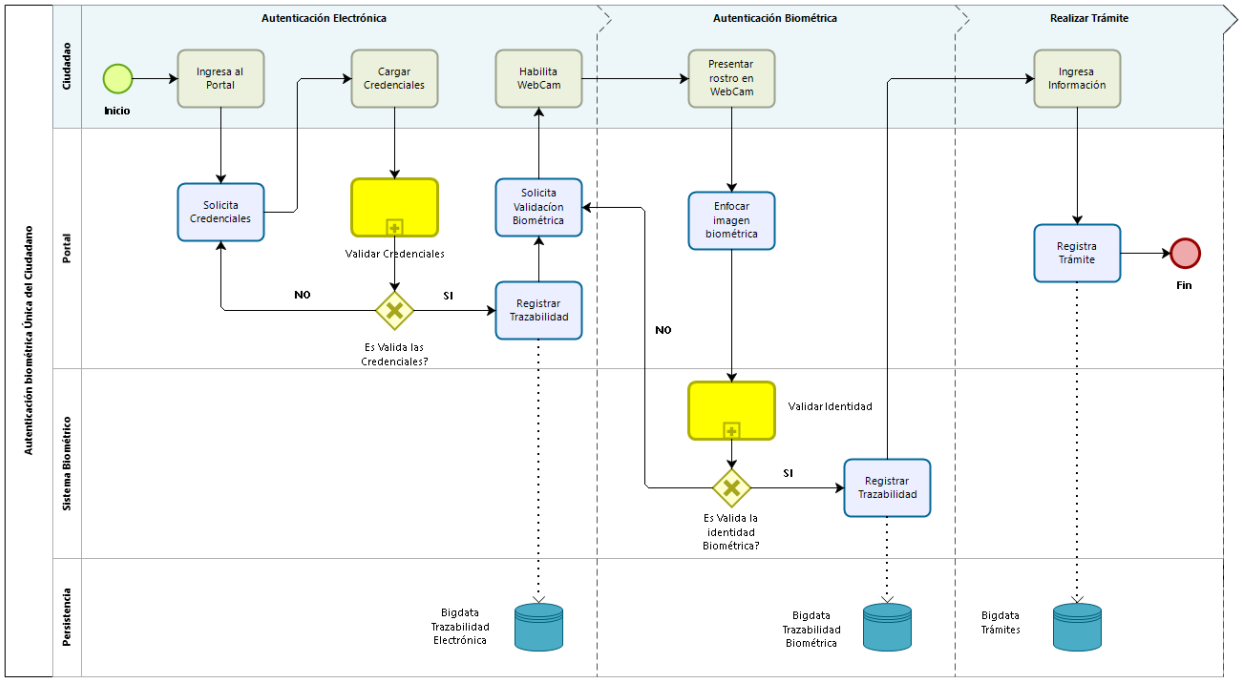


Figura 26, Autenticación biométrica del Ciudadano

Las evaluaciones de los resultados después del entrenamiento se presentan en la Figura 26.

Figura 27, Evaluación del entrenamiento y reconocimiento facial (fuente propia)

4.4 Aspectos de Infraestructura

Para la etapa de entrenamiento se utilizó una Laptop con la siguiente configuración:

Tabla 9, Configuración de infraestructura para realizar pruebas

Equipo de Computo	**Fabricante**	**Procesadores Lógicos**	**Velocidad**	**RAM**	**Disco**
Laptop HP	INTEL Core i5	8 (4 núcleos)	2.4Ghz	8 GB	500GB Solido

4.5 Estrategia de implementación

Se define las capas de la implementación de la solución para soportar el nivel de escalabilidad nacional y lo más importante la seguridad de la información.

4.5.1 Cadena de valor End-to-End de la solución Identificación Biométrica Única

La cadena de valor nos ayuda de forma estrategia a comprender el macro-proceso de la solución, dividiéndolo en tres procesos:

- ✓ **Proceso 1**: Generación del perfil biométrica: Proceso por el cual inicia el enrolamiento utilizando una video cámara.

- ✓ **Proceso 2**: Autenticación biométrica [15]: El ciudadano para utilizar la plataforma digital deberá estar previamente registrado biométricamente.
- ✓ **Proceso 3**: Intercambio de información habilitando la Interoperabilidad: Una vez identificado biométricamente al ciudadano este procederá a realizar cualquier trámite digital disponible y podrá habilitar la interoperabilidad interinstitucional.

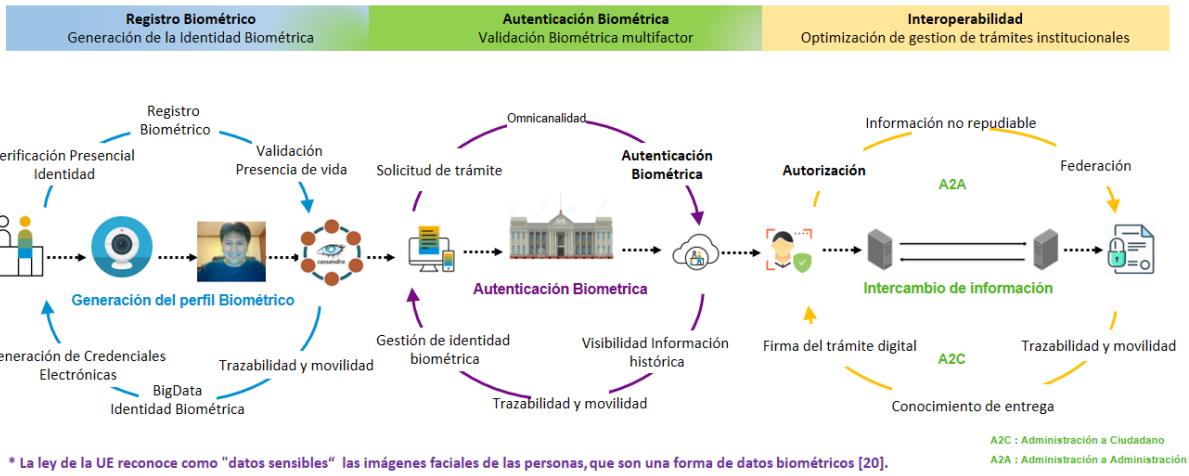


Figura 28, Cadena de Valor End to End para la Identificación Biométrica Única (fuente propia)

El modelo para habilitar la interoperabilidad, está dado por la siguiente propuesta.

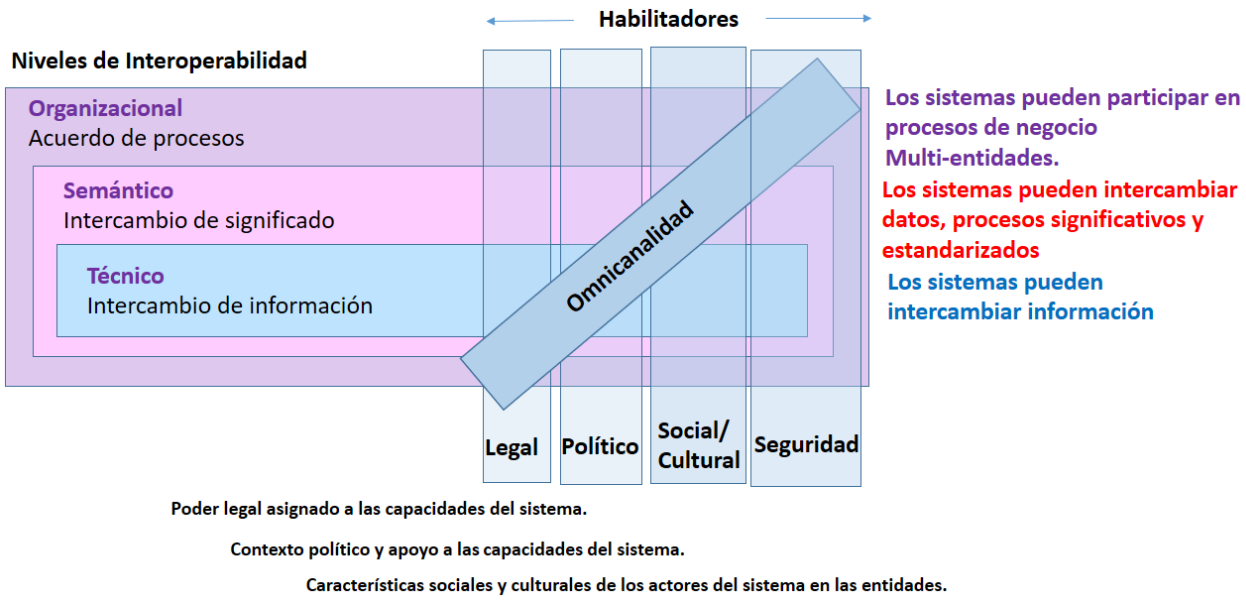


Figura 29, Factores de Influencia en la Estrategia de Interoperabilidad (fuente propia)

4.6 Diseño de implementación

El siguiente esquema muestra la estrategia de implementación considerando elementos de negocio, aplicación, tecnología, seguridad entre otros contextos.

La motivación está relacionada con la oportunidad de crear una solución que ayude a optimizar los recursos económicos y técnicos del estado peruano.

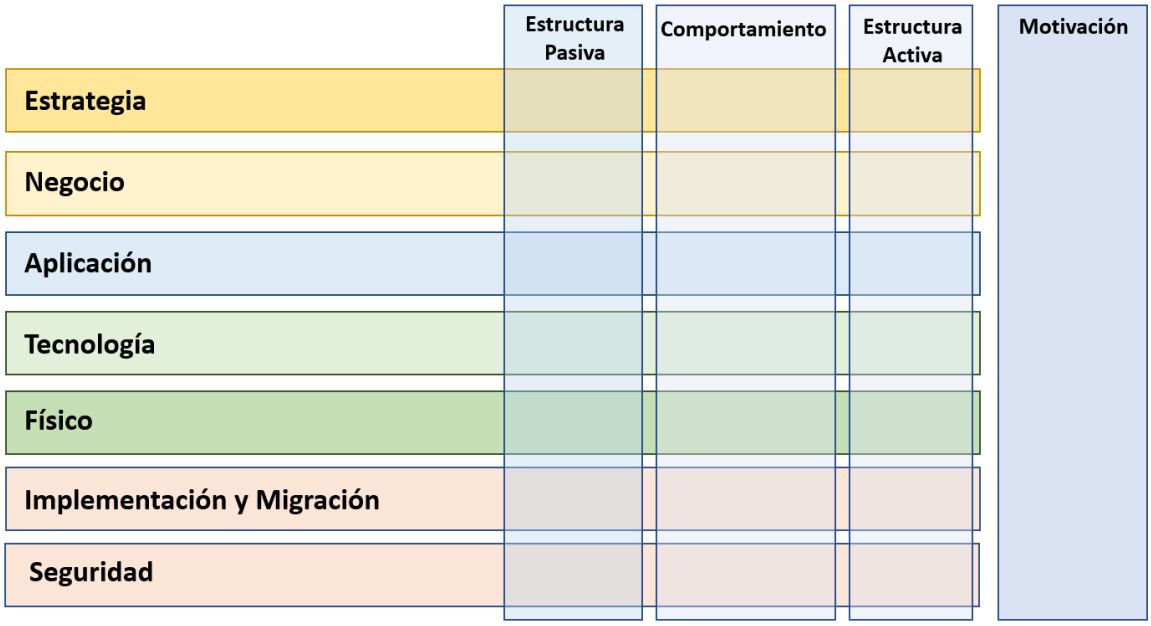


Figura 30, Diseño de implementación (fuente propia)

La siguiente tabla describe el entorno propuesto que serán los ejes que ayudarán en su implementación.

Tabla 10, Diseño de implementación (fuente propia)

Diseño	**Estructura Pasiva**	**Comportamiento**	**Estructura Activa**	**Motivación**
Estrategia	Plataforma digital Única del Estado Peruano	Interoperabilidad	Gob.pe	Ley N° 27444, aprobado con el Decreto Supremo N° 004-2019-JUS

Negocio		✓ Registro de identidad Biométrica. ✓ Autenticación Biométrica.	✓ Ciudadano ✓ Instituciones	Habilitar la interoperabilidad
Aplicación	Identificación biométrica única	En línea	✓ Ciudadano ✓ Instituciones	
Tecnología	**Front-End** ✓Angular 12+ **Back-End** ✓Spring Cloud Netflix ✓Tensorflow **Persistencia** ✓Cassandra ✓PostgreSQL	✓ Ajax ✓ Microservicios ✓ DeepLearning ✓ BigData	Gob.pe	Escalación Nacional y Global
Físico	✓ Servidores en la nube. ✓ Servidores locales.	✓ Virtualización ✓ Contenedores	Gob.pe	Optimización de recursos
Implementación & Migración		Migración de Identidades biométricas RENIEC	Gob.pe	

Seguridad	Autoridad Certificadora	✓ Emisión y Validación de Certificados. ✓ Trasmisión de información vía https.	Gob.pe	Garantizar el intercambio de información y la no repudiación.

4.7 Diseño de Arquitectura

4.7.1 Arquitectura Semántica

La interoperabilidad semántica asegura que el formato preciso y el significado de los datos y la información intercambiados se conserven y se comprendan en los intercambios entre las partes interesadas, en otras palabras, "lo que se envía es lo que se entiende". la interoperabilidad semántica cubre aspectos semánticos y sintácticos:

- ✓ El aspecto semántico se refiere al significado de los elementos de datos y la relación entre ellos. Incluye el desarrollo de vocabularios y esquemas para describir los intercambios de datos, y asegura que los elementos de datos sean entendidos de la misma manera por todas las partes que se comunican.

- ✓ El aspecto sintáctico se refiere a la descripción del formato exacto de la información que se intercambiará en términos de gramática y formato.

Vocabulario Core:

Tabla 11, Entidad de Identificación biométrica única (fuente propia)

Ítem	Identidad Biométrica [supercolumn]	Tipo de Dato	requerido	Descripción
1.	idglobalbiometrico	String [400]	SI	llave única de la identidad Biométrica

2.	primernombre	String [40]	SI	nombres del ciudadano
3.	segundonombre	String [40]	NO	nombres del ciudadano
4.	primerapellido	String [40]	SI	apellido paterno del ciudadano
5.	Segundoapellido	String [40]	NO	apellido materno del ciudadano
6.	dni	String [12]	SI	documento nacional de identidad
7.	pasaporte	String [10]	NO	pasaporte
8.	tipodocumento	String [20]	NO	tipo de documento
9.	nrodocumento	String [40]	NO	número de documento
10.	contrasenia	string[100]	SI	Contraseña (cifrada)

Figura 31, Identidad Biométrica

Se generó en cassandra la siguiente keyspace = biokeyspace, con **replicación** 1, y estrategia **NetworkTopologyStrategy** esta topología es multi data center [12].

Cql

<in>: **CREATE KEYSPACE** biokeyspace **WITH REPLICATION** = {'class' : 'NetworkTopologyStrategy', 'datacenter1' : 1};

<out> : <cassandra.cluster.ResultSet at 0x20d44493160>

Cql: **CREATE TABLE** biokeyspace.trazabilidad (idtrazabilidad uuid **PRIMARY KEY**,idglobalbiometrico text,nombreproceso text,fechaautenticacion timestamp,identidad blob, rostroextraido blob, ciudadano_detectado text ,idclase_detectado int,prescition_detectado float,entrenamiento_h5 text, estatus int)

Se generó la siguiente tabla para la trazabilidad de autenticación:

Tabla 12, Entidad para la trazabilidad de autenticación biométrica única (fuente propia)

Ítem	Trazabilidad [supercolumn]	Tipo de Dato	requerido	Descripción
1.	Idtrazabilidad (PK)	uuid	SI	identificador de la operación
2.	Idglobalbiometrico	Text	NO	llave única de la identidad biométrica del ciudadano
3.	nombreproceso	Text	SI	cualquier trámite requerido
4.	fechaautenticacion	timestamp	SI	fecha de autenticación
5.	identidad	Blob	SI	Imagen del ciudadano autenticado
6.	rostroextraido	Blob	SI	Imagen del rostro extraído del ciudadano autenticado
7.	ciudadano_detectado	Text	SI	Nombre de la clase detectado por la red neuronal, en el momento que el ciudadano está autenticándose.
8.	idclase_detectado	Int	SI	Identificador de clase
9.	prescition_detectado	Float	SI	Precisión de la red neuronal al reconocer la imagen del ciudadano al autenticarse.
10.	entrenamiento_h5	text	SI	Archivo de la red neuronal convolucional entrenada, utilizado para el reconocimiento biométrico del ciudadano.
11.	estatus	int	SI	1. Acceso exitoso 2. No existe en base de datos

Servicios de Integración: Los servicios de negocio están definidos en la siguiente tabla, estos servicios están en un contexto general.

Tabla 13, Servicios utilizados para habilitar la identificación biométrica única (fuente propia)

Ítem	Servicios de Negocio	Verbo	Contexto	Descripción
1.	/RegistroIdentidadBiometrico	POST	interoperabilidad/	Registra las fotos tomadas en el enrolamiento del ciudadano
2.	/APIIdentificacionBiometricaUnicaCNN	POST	interoperabilidad/	Validación de la identidad biométrica de la del ciudadano
3.	/APIEntrenamientoCNN	POST	interoperabilidad/	Entrena las redes Neuronales, en función al tipo de modelo VGG16, VGG19, RESTNET50
4.	/API-BigDataCluster	POST	Trazabilidad	Registro de trazabilidad de autenticación e identificación biométrica
5.	/API-Matrizconfusion	POST	Evaluación de los Modelos VGG16,	Evaluación del desempeño del entrenamiento

			VGG19, RESTNET50	de las redes neuronales convolucionales utilizando la matriz de confusión para cada modelos

Generación de certificados de seguridad para la publicación de la plataforma de Identidad Biométrica Única:

Tabla 14, Generación de certificado SSL de seguridad para la plataforma de Identidad Biométrica

Ítem	**Configuración y comandos**	**Resultado**
1.	openssl genrsa -des3 -out ca.key 4096	Ca.key
2.	openssl req -new -x509 -days 3650 -key ca.key -out ca.crt Enter pass phrase for ca.key: You are about to be asked to enter information that will be incorporated into your certificate request. What you are about to enter is what is called a Distinguished Name or a DN. There are quite a few fields but you can leave some blank For some fields there will be a default value,	Ca.crt

	If you enter '.', the field will be left blank. --- Country Name (2 letter code) [AU]:PE State or Province Name (full name) [Some-State]:LIMA Locality Name (eg, city) []:LIMA Organization Name (eg, company) [Internet Widgits Pty Ltd]: IBU Organizational Unit Name (eg, section) []:TI Common Name (e.g. server FQDN or YOUR name) []:IdentidadBiometricaUnica Email Address []:identidaddigital@pcm.gob.pe	
3.	openssl rsa -in ca.key -text > ca.pem	ca.pem
4.	openssl x509 -inform PEM -in ca.crt >> ca.pem	ca.pem
5.	openssl genrsa -des3 -out user.key 4096	User.key
6.	openssl req -x509 -nodes -days 365 -newkey rsa:2048 -keyout private-nginx-selfsigned.key -out nginx-selfsigned.crt	Creación del certificado SSL utilizando OpenSSL, para asegurar el NGINX utilizando HTTPS

4.7.2 Arquitectura Tecnológica

La solución está definida para los siguientes componentes y tecnologías, como para la parte front-end utilizando el framework de angular, así como para el back-end utilizando los micro-servicios desarrollados en Spring Boot y su integración de datos en PostgreSQL y Cassandra.

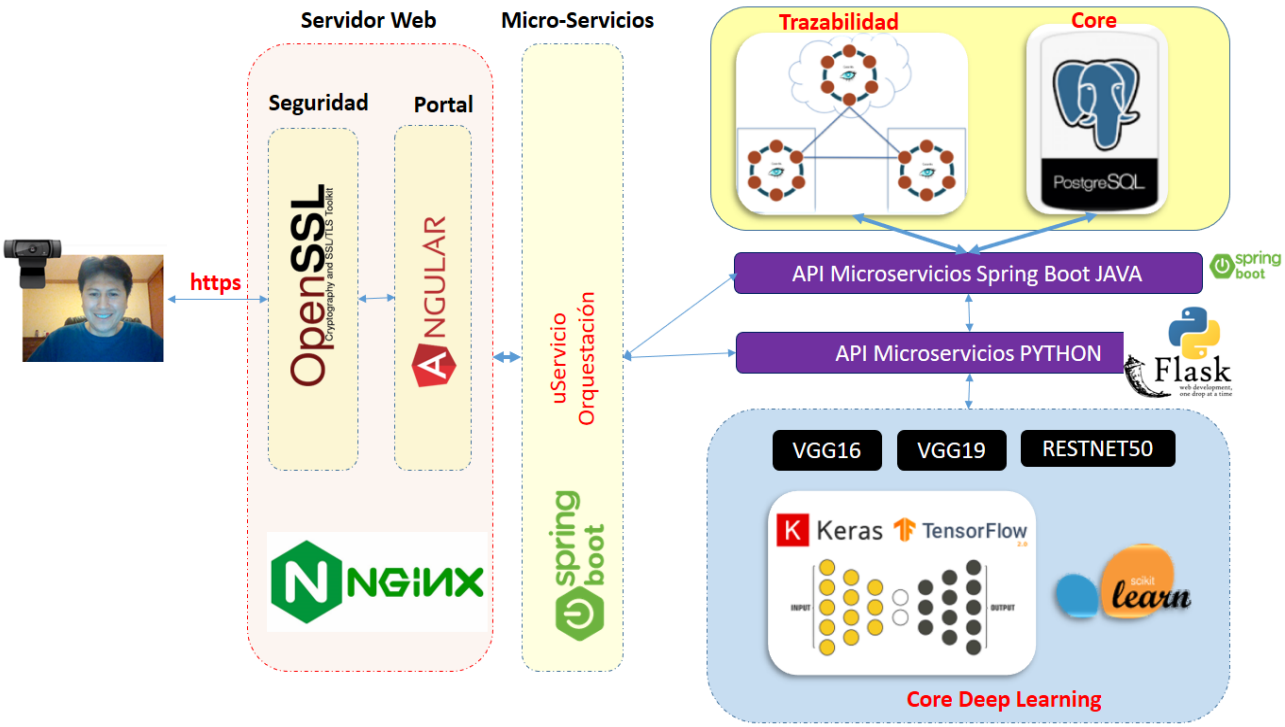


Figura 32, Modelo conceptual de la arquitectura tecnológica de solución (fuente propia)

Los componentes tecnológicos utilizados para la implementación de la solución en la siguiente tabla con sus respectivas descripciones.

Tabla 15, Componentes tecnológicos utilizandos

Ítem	Componente Tecnológico	Descripción
1.	NGINX	Servidor web estático, es un servidor web/proxy inverso ligero de alto rendimiento. Originalmente, Nginx fue desarrollado para satisfacer las necesidades de varios sitios web de Rambler que recibían unas 500 millones de peticiones al día en septiembre de 2008.(https://www.linuxjournal.com/article/10108)
2.	OpenSSL	Openssl es una api que proporciona un entorno adecuado para encriptar los datos enviados a otra computadora dentro de una red y a su vez des

		encriptarlos adecuadamente por el receptor, evitando así, el acceso a la información por intrusos con la utilización de sniffer.
3.	Angular	Angular es un framework para aplicaciones web desarrollado en TypeScript, de código abierto, mantenido por Google, que se utiliza para crear y mantener aplicaciones web de una sola página.
4.	Spring Boot	Spring Boot es un micromarco de código abierto mantenido por una empresa llamada Pivotal. Proporciona a los desarrolladores de Java una plataforma para empezar con una aplicación Spring autoconfigurable de grado de producción.
5.	Flask Python	Flask es un framework minimalista escrito en Python que permite crear aplicaciones web rápidamente y con un mínimo número de líneas de código. Está basado en la especificación WSGI de Werkzeug y el motor de templates Jinja2 y tiene una licencia BSD.
6.	Keras	Keras es una biblioteca de Redes Neuronales de Código Abierto escrita en Python. Es capaz de ejecutarse sobre TensorFlow, Microsoft Cognitive Toolkit o Theano.Está especialmente diseñada para posibilitar la experimentación en más o menos poco tiempo con redes de Aprendizaje Profundo.
7.	Tensorflow	TensorFlow es una biblioteca de código abierto para aprendizaje automático a través de un rango de tareas, y desarrollado por Google para satisfacer sus necesidades de sistemas capaces de construir y entrenar redes neuronales para detectar y descifrar

		patrones y correlaciones, análogos al aprendizaje y razonamiento usados por los humanos[10].
8.	Scikit-learn	La librería scikit-learn, también llamada sklearn, es un conjunto de rutinas escritas en Python para hacer análisis predictivo, que incluyen clasificadores, algoritmos de clusterización, etc. Está basada en NumPy, SciPy y matplotlib, de forma que es fácil reaprovechar el código que use estas librerías, este framework se utiliza para la evaluación del desempeño de los modelos utilizando confusion_matrix.
9.	Cassandra	Apache Cassandra se trata de un software NoSQL distribuido y basado en un modelo de almacenamiento de «clave-valor», de código abierto que está escrita en Java. Permite grandes volúmenes de datos en forma distribuida. Por ejemplo, lo usa Twitter para su plataforma.
10.	PostgreSQL	PostgreSQL, también llamado Postgres, es un sistema de gestión de bases de datos relacional orientado a objetos y de código abierto, publicado bajo la licencia PostgreSQL, similar a la BSD o la MIT

4.7.3 Arquitectura de Solución

La solución está compuesta con una plataforma que ayudara a la evaluación los modelos de entrenamiento de las redes neuronales convolucionales VGG16, VGG19 y RESTNET50.

El enrolamiento biométrico parte de una cámara web, que captura la imagen del rostro, posteriormente utilizando visión artificial es identificado el rostro y posteriormente es extraída utilizando el método de Haar Cascade [1] de OpenCV, y solo la parte del rostro de la imagen preparando así para su posterior entrenamiento utilizando modelos VGG16, VGG19 y RESTNET50.

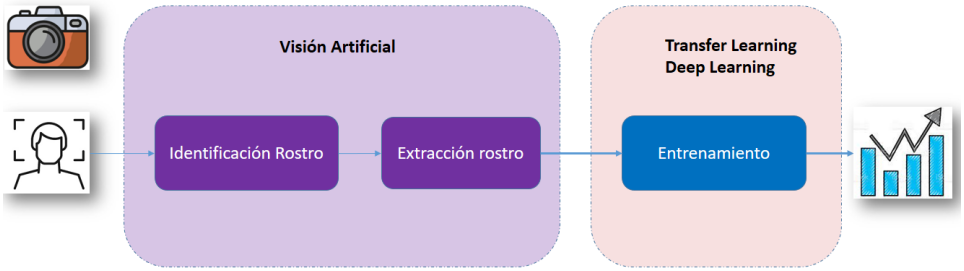


Figura 33, Modelo de enrolamiento biométrico

Arquitectura de los modelos de entrenamiento:

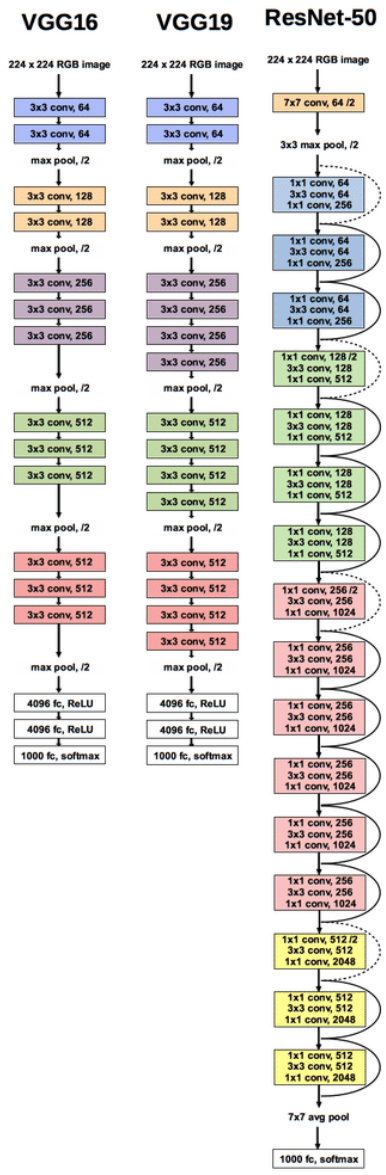


Figura 34, Arquitectura de Modelos VGG16, VGG19, RestNet50

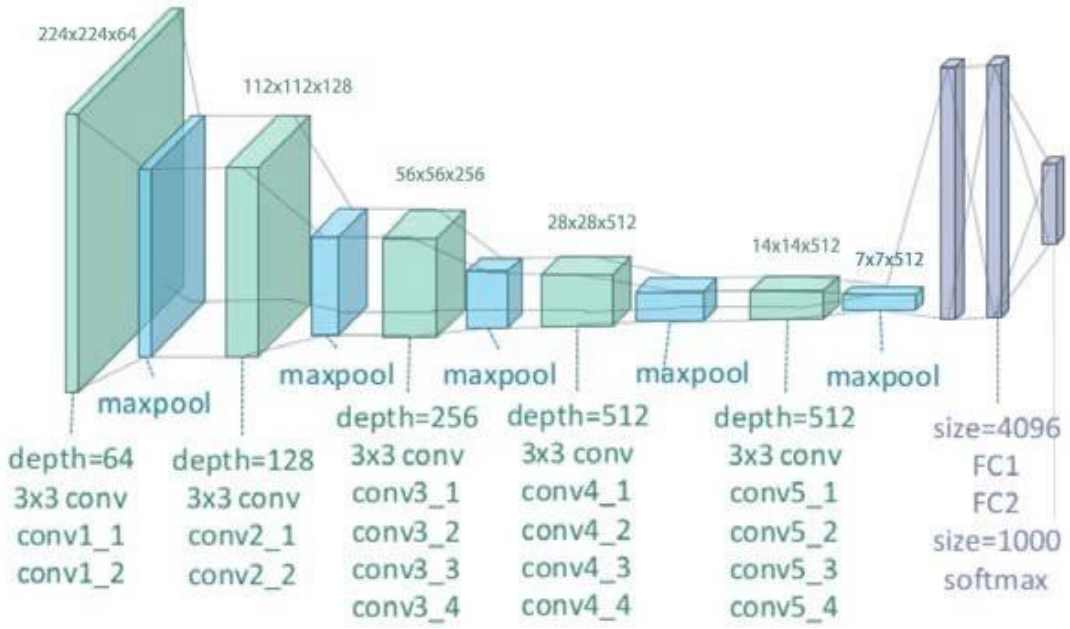


Figura 35, Arquitectura VGG19

La siguiente imagen representa la evaluación biométrica, este proceso sucede cuando el ciudadano esta por autenticarse para posteriormente realizar un trámite de gobierno.

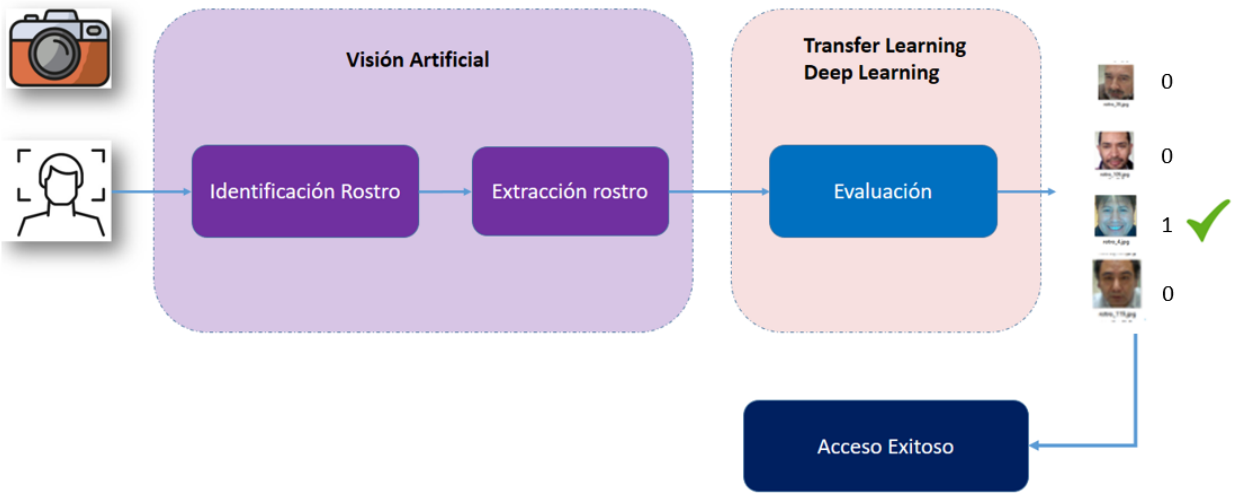


Figura 36, Evaluación Biométrica

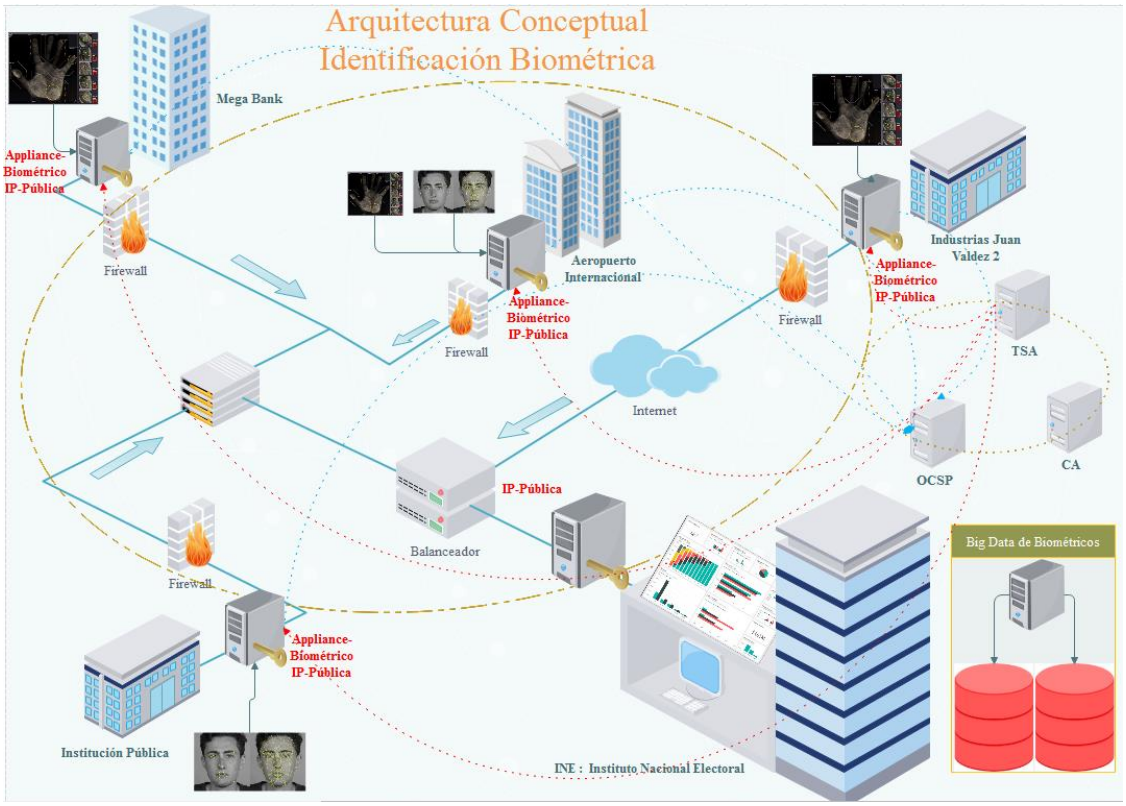


Figura 37, Arquitectura Conceptual de Identificación Biométricos Única (Fuente Propia)

Los componentes de la arquitectura, principalmente están determinados por herramientas de seguridad, como son:

- Autoridad Certificadora (CA): Es una entidad de confianza, responsable de emitir y revocar los certificados, utilizando en ellos la firma electrónica, para lo cual se emplea la criptografía de clave pública (PKI).
- Protocolo de comprobación del Estado de un Certificado En línea (OCSP): Es s un método para determinar el estado de vigencia de un certificado digital X.509 usando otros medios que no sean el uso de CRL (Listas de Revocación de Certificados). Este protocolo se describe en el RFC 6960[27] y está en el registro de estándares de Internet.
- Autoridad de sellado de tiempo (TSA): Es un prestador de servicios de certificación que proporciona certeza sobre la preexistencia de determinados documentos electrónicos a un momento dado, cuya indicación temporal junto con el hash del documento se firma por la

[27] X.509 Internet Public Key Infrastructure, Online Certificate Status Protocol – OCSP, https://datatracker.ietf.org/doc/html/rfc6960

Autoridad de sellado de tiempo. Este sellado de tiempo es uno de los servicios más importantes en relación con la firma electrónica.

- Appliance Biométrico: Es el componente de autenticación biométrica única, que deberá coexistir en cada una de las instituciones y organismos del estado, con el objetivo de identificar la veracidad y autenticidad de los ciudadanos.

  Cada appliance es un componente de interoperabilidad (Security Server), basado en la arquitectura X-ROAD.
- BigData Biométricos: Es la base de datos de biométricos donde se almacenarán todas las imágenes de los ciudadanos, así como toda la trazabilidad de identificación y autenticación en la plataforma de identificación biométrica única.
- Institucional nacional Electoral (INE en México/ RENIE en Perú): Organismo encargado del enrolamiento biométrico de los ciudadanos, generando su carnet de identificación y de elector.

4.8 Comunicaciones Segura

La seguridad es un factor importante sobre todo si se está accediendo a validar identidades biométricas, información sumamente privada y confidencial.

Una forma de resolver este tipo de problemas es habilitando certificados de autenticación para cada ciudadano, donde el servidor validara si el certificado de autenticación es válido, no está caducado y sobre todo que se haya emitido por la Autoridad Certificadora de gob.pe (Plataforma Digital Única del Estado Peruano).

HTTPS, es el canal seguro para la comunicación de los datos biométricos del ciudadano con el servidor.

4.9 Escalamiento horizontal

Existen 35 municipalidades en lima y en el resto del país 209, solamente en municipalidades, y más cantidades de instituciones de salud, de comisarias, educativas, agricultura, minería y otros muchos más.

Los ciudadanos pueden realizar de forma simultanea tramites en cualquier institución de gobierno, en cualquier parte y en cualquier momento, bajo estos requisitos es importante que la plataforma

de interoperabilidad peruana pueda tener las capacidades de poder atender no solo todos los tramites de todas las instituciones, sino que también tenga la capacidad de poder autenticar e identificar biométricamente al ciudadano en cada uno de los tramites.

CAPITULO 5

RESULTADOS, PRUEBAS Y EVALUACIÓN

5.1 Módulos del Sistema

La aplicación cuenta con los siguientes módulos:

1. **Módulo de Registro**: Este módulo habilita la una cámara web, para poder capturar las imágenes y la asociación de la clase, ejemplo su DNI, nro. Pasaporte, Otros documentos.
2. **Módulo de Lista de Rostros Registrados**: Este módulo, se puede realizar la búsqueda por identificador biométrico, el sistema visualizara la lista de las imágenes registradas, con tres vistas, la primera la foto capturada, la segunda vista la foto capturada y remarcado en un rectángulo el rostro, y la tercera vista el rostro extraído.

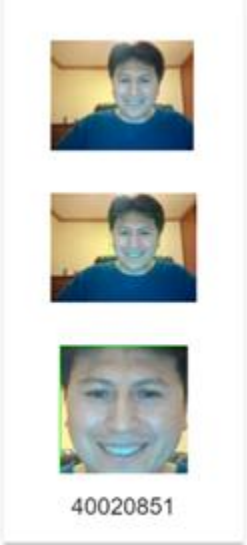


Figura 38, Ejemplo del módulo listar rostros registrados(fuente propia)

3. **Módulo Entrenamiento**: Este módulo realiza el entrenamiento tomando como parámetros el modelo de red neuronal convolucional VGG16, VGG19 y RESTNET50, así como el número de épocas los cuales podrá realizar en el proceso de entrenamiento, visualizando al final del entrenamiento los resultados de la evaluación, precisión y % de perdida realizada con los parámetros proporcionados.

4. **Módulo de Clases Entrenadas**: Realiza una consulta a partir del identificador de entrenamiento y visualiza las imágenes de las clases entrenadas.
5. **Módulo de Identificación Biométrica**: Realiza una prueba de validación de identificación biométrica, habilitando la cámara web y en tiempo real validando la persona su identificación, registrando así la trazabilidad.
6. **Módulo de Trazabilidad de Autenticación**: este módulo visualiza a partir de filtros como el identificador biométrico, rango de fechas, así como el número de registros que se requieren visualizar, este módulo es alimentado cada vez que se realiza una identificación biométrica única.

La siguiente figura muestra la vista principal del sistema, identificando el menú de cada módulo.

Figura 39, Módulos del sistema de identificación biométrica única

La siguiente tabla presenta la descripción de los módulos implementados.

Tabla 16, Lista de componentes implementados en la solución

Ítem	Modulo	Descripción
1.	Registrar	El módulo habilita una cámara web y captura la foto de la persona, asociándole el identificador de clase/clasificación, en esta etapa del enrolamiento interviene los algoritmos visión por computadora, procesando la imagen y extrayendo únicamente el rostro de la foto.
2.	Listar Rostros	Buscar a partir del identificador de clase la foto del rostro de la persona/ciudadano, visualizando los diferentes

		momentos, la foto original, y su respectiva detección del rostro en la imagen y su extracción.
3.	Entrenamiento	El módulo tiene como propósito de lanzar el proceso de entrenamiento de las redes neuronales convolucionales, utilizando como parámetros de entrada el número de épocas, y el respectivo modelo VGG16, VGG19, RESTNET50.
4.	Lista de Clases	Una vez terminado el entrenamiento la red neuronal, internamente el sistema almacena en base de datos la clase identificada asociando su respectiva imagen del rostro de la persona.
5.	Evaluación de Modelos	El módulo de evaluación, evalúa los desempeños de los modelos entrenados utilizando la matriz de confusión, este módulo tiene como entrada los archivos uuid_modelo.h5 generados en la fase de entrenamiento, generando y visualizando la distribución de los registros en términos de clases reales y clases pronosticadas como se muestran en la figura 49.
6.	Identificación de Rostros	El módulo habilita una cámara web, y cada 2 seg. Captura un Frame que es enviado primero al proceso de extracción de rostros utilizando algoritmos de visión por computadora(OpenCV), para que la salida de este proceso pase directamente a la red neuronal entrenada para su respectiva predicción, identificando así la clase correspondiente.
7.	Trazabilidad	El módulo de trazabilidad, ayudara a registrar cada evento de intento de identificación biométrico, este módulo registra, fecha, hora, modulo, precisión de la red neuronal en la predicción, así como el identificado clasificado entre otros campos.

Para las pruebas realizadas se utilizó la siguiente configuración de imágenes para los tres modelos teniendo en cuenta la arquitectura de cada modelo presentado en la Figura 33 del capítulo anterior.

La siguiente imagen muestra la vista de las 9 clases utilizada para el entrenamiento, y se consideraron para cada clase 300 imágenes, dando un 80% para el entrenamiento y el 20% para la evaluación.

Cada clase es capturada por medio de una video cámara, y cada imagen capturada pasa por un proceso de identificación de rostro y su respectiva extracción a 224x224 que es lo que requieren los modelos como entrada.

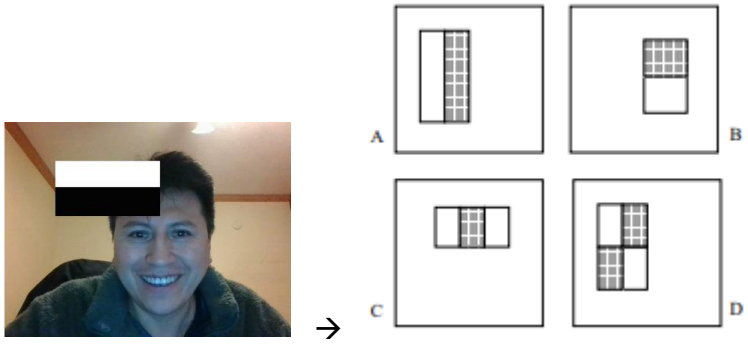


Figura 40, Recorrido horizontal de las etapas para el cálculo del haar [13].

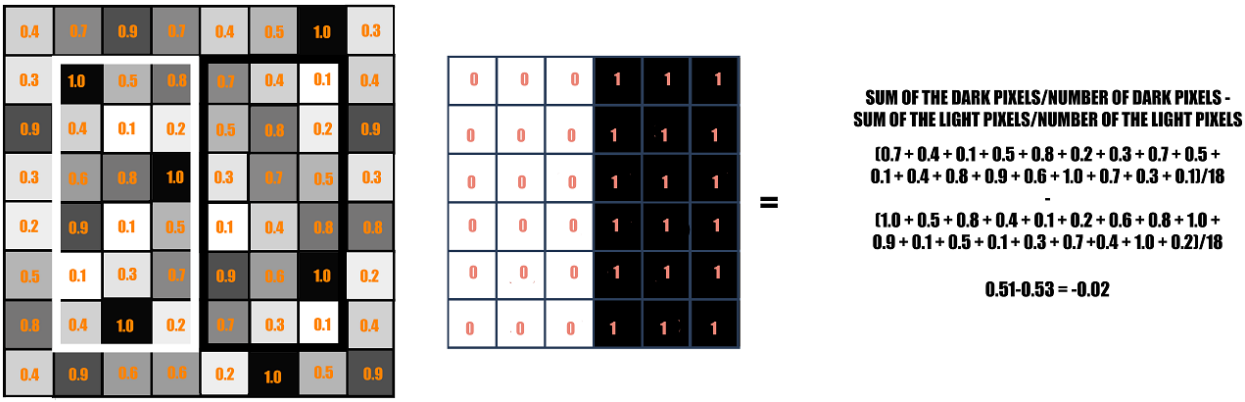


Figura 41, Operación para el cálculo de haar, la existencia de contorno

El ejemplo presentado de características de los rectángulos mostradas en relación con la ventana de detección que los rodea. La suma de los píxeles que se encuentran dentro de los rectángulos blancos se resta de la suma de los píxeles de los rectángulos grises. Las características de dos rectángulos se muestran en (A) y (B). La figura (C) muestra una característica de tres rectángulos, y (D) una característica de cuatro rectángulos.

Utilizando el algoritmo y las librerías de OpenCV [13],[16], la siguiente línea de código fuente en python 3+ identifica y remarca con un rectángulo específicamente rostros en una imagen.

- grayscale_image = cv2.cvtColor(original_image, cv.COLOR_BGR2GRAY)

- face_cascade = cv2.**CascadeClassifier**('haarcascade_frontalface_alt.xml')[1]

- detected_faces = face_cascade.**detectMultiScale**(grayscale_image)
- for (column, row, width, height) in detected_faces:
  cv2.rectangle(original_image,(column, row),(column + width, row + height),(0, 255, 0),4)
- plt.imshow(cv2.cvtColor(original_image, cv2.COLOR_BGR2RGB))
- plt.axis("off")
- plt.show()

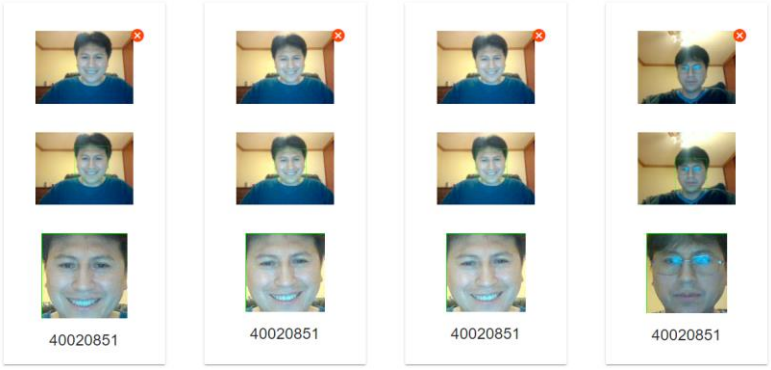


Figura 42, Enrolamiento, identificación y extracción de rostros utilizando OpenCV

El objetivo aquí es averiguar la suma de todos los píxeles de la imagen que se encuentran en el área más oscura de la función haar y la suma de todos los píxeles de la imagen que se encuentran en el área más clara de la función haar. y luego se descubre su diferencia. Si la imagen tiene un borde que separa los píxeles oscuros a la derecha y los claros a la izquierda, entonces el valor de haar estará más cerca de 1. Eso significa que decimos que hay un borde detectado si el valor de haar está más cerca de 1. En la Figura 42, no hay borde ya que el valor de haar está lejos de 1.

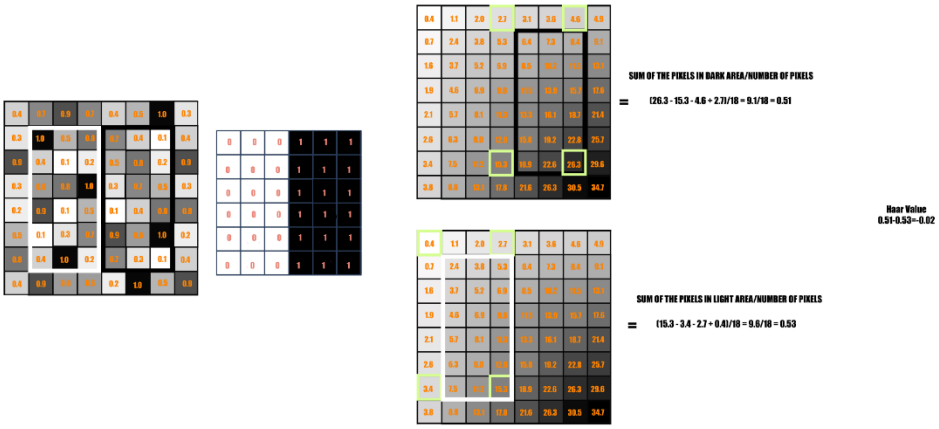


Figura 43, la imagen integral para calcular el haar

La siguiente Figura representa la muestra de 9 clases utilizado para las respectivas pruebas y validación de los modelos según sea el caso, se consideró para cada clase 300 imágenes con una configuración de 224x224 pixeles recomendados según la bibliografía [2].

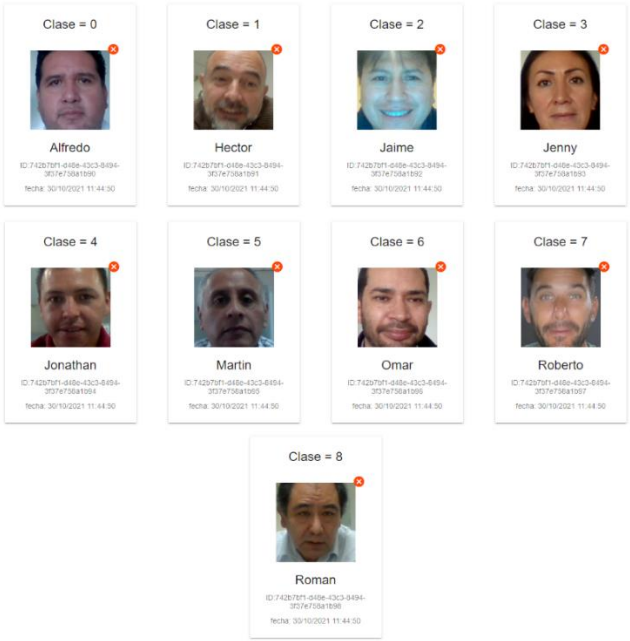


Figura 44, lista de las 9 clases utilizadas para el entrenamiento de los diferentes modelos

Visualizando las características por bloques de la red neuronal convolucional entrenada.

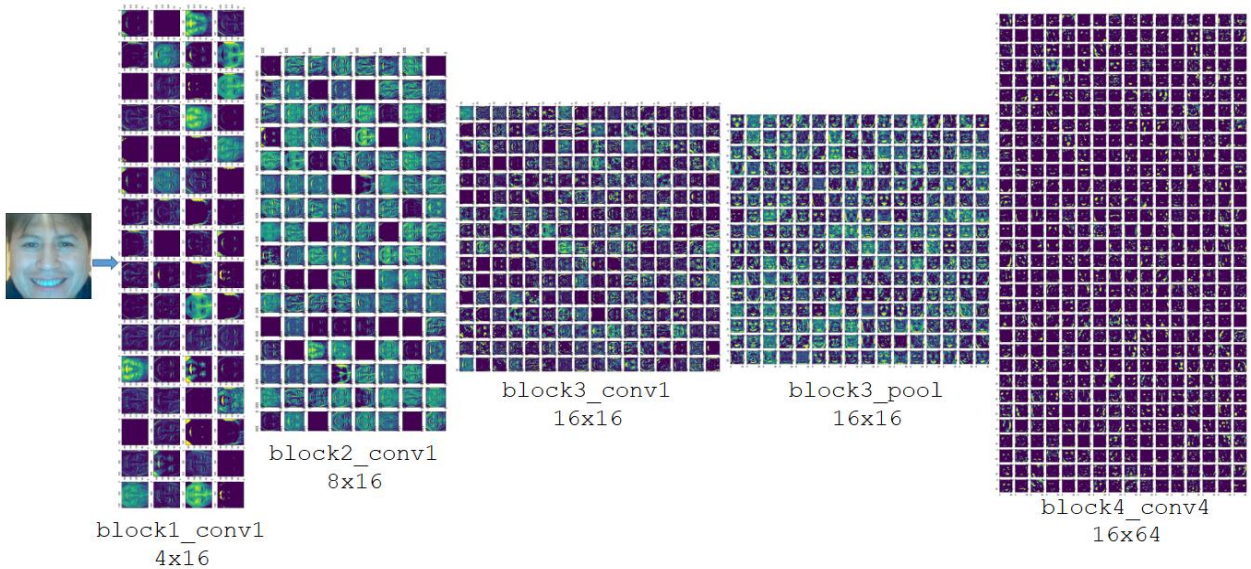


Figura 45, Mapa de características por bloques de la red neuronal convolucional entrenada

Ejemplo de los mapas de características generados por la red neuronal VGG16, después de su entrenamiento.

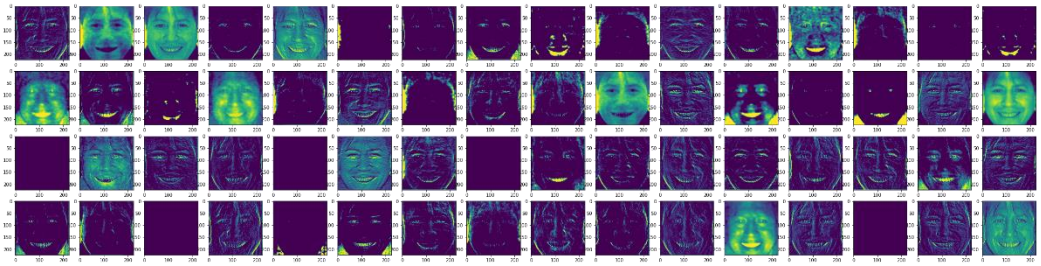


Figura 46, Visualización de características del primer bloque utilizando VGG16 (block1_conv1 4x16)

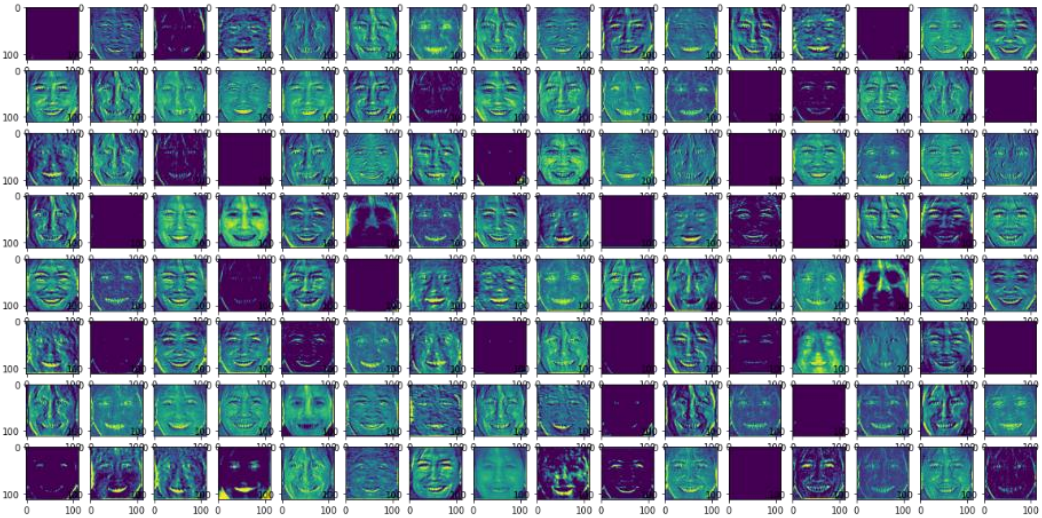


Figura 47, Visualización de características del segundo bloque utilizando VGG16 (block2_conv1 8x16)

5.2 Pruebas y Resultados

La siguiente imagen muestra la precisión y la perdida por modelo, así como un gráfico comparativo de cada entrenamiento y evaluación para cada modelo.

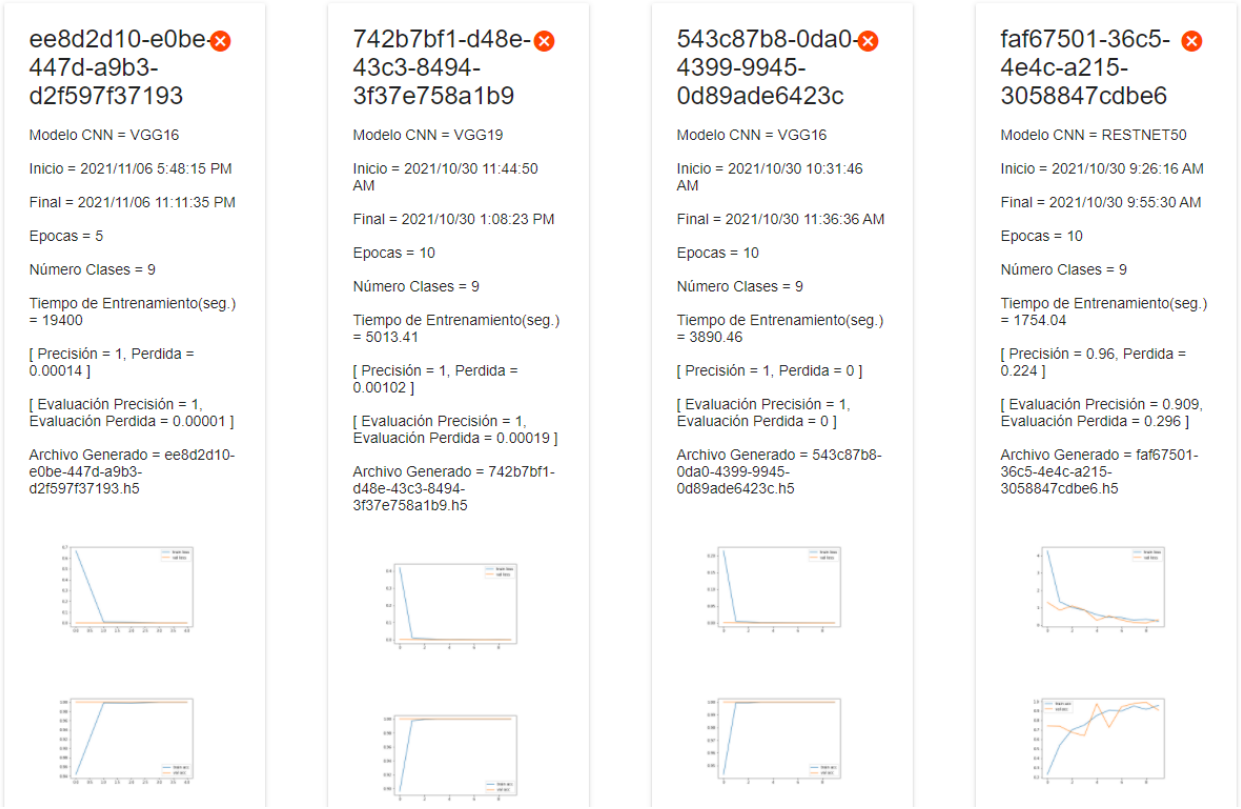


Figura 48, resultado y evaluación de los modelos VGG16, VGG19, RESTNET50 y VGG16+3 (fuente propia)

Los resultados del entrenamiento y de la evaluación en la siguiente tabla.

Tabla 17, Tabla de comparación de resultados para los modelos VGG16, VGG19 y RESTNET50

Evaluación CNN	**VGG16**	**VGG19**	**RESTNET50**	**VGG16+3**
Numero de épocas	10	10	10	5
Numero de clases	9	9	9	9
Tiempo Entrenamiento	3890 seg. (1.08 Hora)	5013 seg. (1.39 Hora)	1754 seg. (29.23 Minutos)	19400 seg. (5.39 Hora)
# de imágenes por clase	300	300	300	300
# de imágenes para entrenamiento	240	240	240	240

# de imágenes para entrenamiento	60	60	60	60
Precisión	1	1	0.9600	1
Perdida	0	0.00102	0.2240	0.00014
Evaluación Precisión	1	1	0.9090	1
Evaluación Perdida	0	0.00019	0.2960	0.00001

El proceso de evaluación de los modelos se utiliza una estrategia de las mismas condiciones, así como las mismas entradas de resolución para cada imagen.

Los parámetros de entrada son los archivos de salida uuid_modelo.h5 generados en el proceso de entrenamiento, los resultados del desempeño del entrenamiento en la siguiente imagen, vista en la siguiente figura del análisis del desempeño de los modelos entrenados utilizando la matriz de confusión.

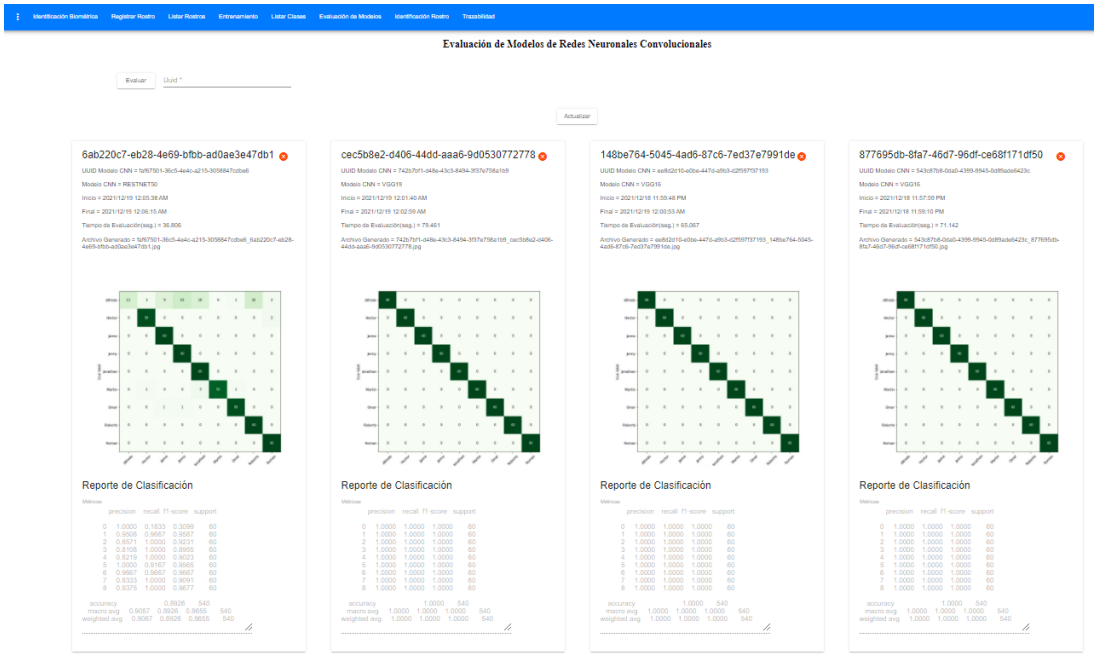


Figura 49, Evaluación del desempeño de modelos entrenados VGG16, VGG19 y RESTNET50

Para la identificación biométrica, se considera las mismas condiciones para cada modelo, tomando como entrada un frame cada 2 segundos capturado a partir de una cámara web, validando y utilizando la red neuronal entrenada.

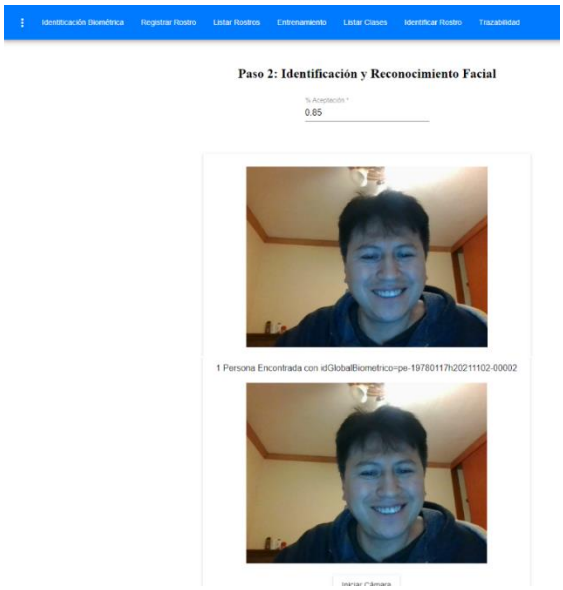


Figura 50, Evaluación de la red neuronal entrenada

Para realizar la identificación biométrica facial se habilita la APIIdentificacionBiometricaUnicaCNN, generando siguiente trazabilidad y la identificación biométrica de la clase 2, como se muestra en la figura 48.

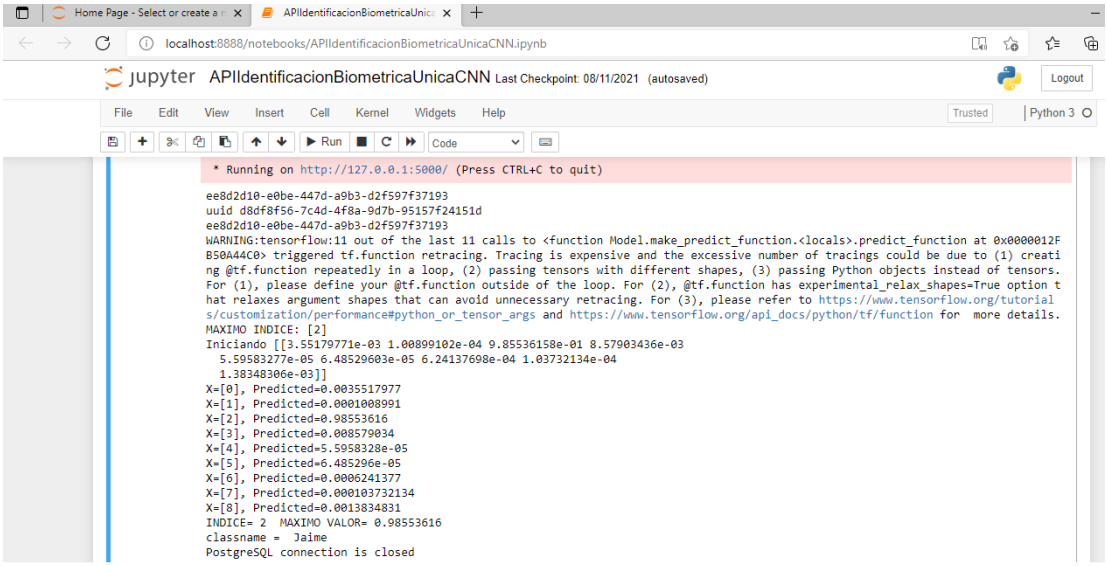


Figura 51, Logs generado en el proceso de reconocimiento facial

La trazabilidad de autenticación biométrica única, se registra en la base de datos NoSQL columnar Cassandra, con los siguientes datos, Id global Biométrico, id de trazabilidad, el proceso donde se está autenticando, fecha de autenticación, estatus de autenticación y su respectiva foto de su rostro almacenada.

Ítem	Id Global Biométrico	Id Trazabilidad	Proceso	Fecha de Autenticación	Nombre de clase detectado	Prescition detectado	ID Clase detectado	Entrenamiento h5	Estatus	Ciudadano
1	pe-19780117h20211102-00002	605a96a6-b1e2-4ba8-bb65-347e2721c334	factura electrónica	2021/11/07 10:00:27 PM	Jaime	0.7848527	2	ee8d2d10-e0be-447d-a9b3-d2f597f37193.h5	0	
2	pe-19780117h20211102-00002	efbb7294-5964-49d4-91d6-13f534e94c2f	factura electrónica	2021/11/07 10:00:18 PM	Jaime	0.9999964	2	ee8d2d10-e0be-447d-a9b3-d2f597f37193.h5	1	
3	pe-19780117h20211102-00002	ab9e14f1-f3d9-457e-a105-dadb06dde516	factura electrónica	2021/11/07 10:00:06 PM	Jaime	0.7685211	2	ee8d2d10-e0be-447d-a9b3-d2f597f37193.h5	0	
4	pe-19780117h20211102-00002	18f80bdf-fdf8-4d97-b41f-2e0fa302bcc1	factura electrónica	2021/11/07 10:00:14 PM	Jaime	0.9999957	2	ee8d2d10-e0be-447d-a9b3-d2f597f37193.h5	1	
5	pe-19780117h20211102-00002	64ff663a-7428-4887-b863-e3951857917f	factura electrónica	2021/11/07 10:00:18 PM	Jaime	0.99999857	2	ee8d2d10-e0be-447d-a9b3-d2f597f37193.h5	1	

Figura 52, trazabilidad de autenticación biométrica (fuente propia)

5.3 Evaluación de Resultados

Se realizaron las evaluaciones de desempeño de los modelos VGG16, VGG19, y RESTNET50 utilizando la matriz de confusión para cada uno de los modelos, y los resultados del análisis en los siguientes puntos.

5.1.1 Evaluación VGG16

- **Identificador**: ee8d2d10-e0be-447d-a9b3-d2f597f37193
- Tamaño de archivo generado del entrenamiento: ee8d2d10-e0be-447d-a9b3-d2f597f37193.h5 (207 MB)
- Numero de clases: 9
- Cantidad de imágenes a evaluar por clase: 60 imágenes por clase
- Total, imágenes evaluadas:540
- Tiempo de entrenamiento: 19400 segundos (5.389 horas).

- Tiempo de Evaluación = 65.067 segundos.

Tabla 18, Tabla de evaluación del modelo VGG16+3

	precision	recall	f1-score	support
0	1.0000	1.0000	1.0000	60
1	1.0000	1.0000	1.0000	60
2	1.0000	1.0000	1.0000	60
3	1.0000	1.0000	1.0000	60
4	1.0000	1.0000	1.0000	60
5	1.0000	1.0000	1.0000	60
6	1.0000	1.0000	1.0000	60
7	1.0000	1.0000	1.0000	60
8	1.0000	1.0000	1.0000	60
accuracy			1.0000	540
macro avg	1.0000	1.0000	1.0000	540
weighted avg	1.0000	1.0000	1.0000	540

La siguiente figura muestra los resultados de la evaluación del modelo VGG16+3, utilizando la matriz de confusión.

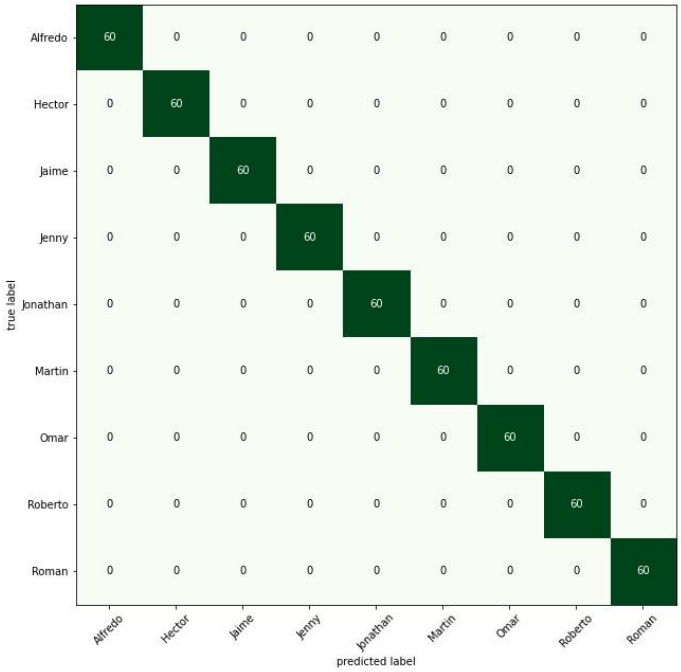


Figura 53, Evaluación matriz de confusión para VGG16+3

5.1.2 Evaluación VGG19

- Identificador: 742b7bf1-d48e-43c3-8494-3f37e758a1b9
- Tamaño de archivo generado del entrenamiento: 742b7bf1-d48e-43c3-8494-3f37e758a1b9.h5 (81 MB)
- Numero de clases: 9
- Cantidad de imágenes a evaluar por clase: 60 imágenes por clase
- Total, imágenes evaluadas:540
- Tiempo de entrenamiento: 5013.41 segundos (1.39 horas).
- Tiempo de Evaluación = 79.461 segundos.

Tabla 19, Tabla de evaluación del modelo VGG19

	precision	recall	f1-score	support
0	1.0000	1.0000	1.0000	60
1	1.0000	1.0000	1.0000	60
2	1.0000	1.0000	1.0000	60
3	1.0000	1.0000	1.0000	60
4	1.0000	1.0000	1.0000	60
5	1.0000	1.0000	1.0000	60
6	1.0000	1.0000	1.0000	60
7	1.0000	1.0000	1.0000	60
8	1.0000	1.0000	1.0000	60
accuracy			1.0000	540
macro avg	1.0000	1.0000	1.0000	540
weighted avg	1.0000	1.0000	1.0000	540

La siguiente figura muestra los resultados de la evaluación del modelo VGG19, utilizando la matriz de confusión.

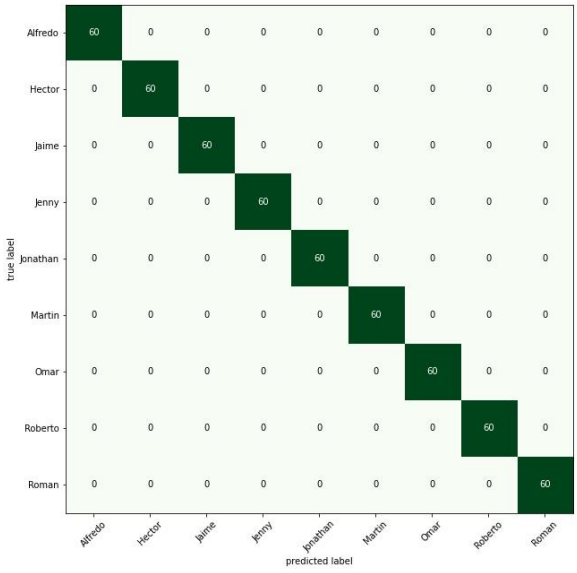


Figura 54, Evaluación matriz de confusión para VGG19

5.1.3 Evaluación RESTNET50

- Identificador: faf67501-36c5-4e4c-a215-3058847cdbe6
- Tamaño de archivo generado del entrenamiento: faf67501-36c5-4e4c-a215-3058847cdbe6.h5 (103 MB)

- Numero de clases: 9
- Cantidad de imágenes a evaluar por clase: 60 imágenes por clase
- Total, imágenes evaluadas:540
- Tiempo de entrenamiento: 1754.04 segundos (0.49 horas).
- Tiempo de Evaluación = 36.806 segundos.

Tabla 20, Tabla de evaluación del modelo RESTNET50

	precision	recall	f1-score	support
0	1.0000	0.1833	0.3099	60
1	0.9508	0.9667	0.9587	60
2	0.8571	1.0000	0.9231	60
3	0.8108	1.0000	0.8955	60
4	0.8219	1.0000	0.9023	60
5	1.0000	0.9167	0.9565	60
6	0.9667	0.9667	0.9667	60
7	0.8333	1.0000	0.9091	60
8	0.9375	1.0000	0.9677	60
accuracy			0.8926	540
macro avg	0.9087	0.8926	0.8655	540
weighted avg	0.9087	0.8926	0.8655	540

La siguiente figura muestra los resultados de la evaluación del modelo RESTNET50, utilizando la matriz de confusión.

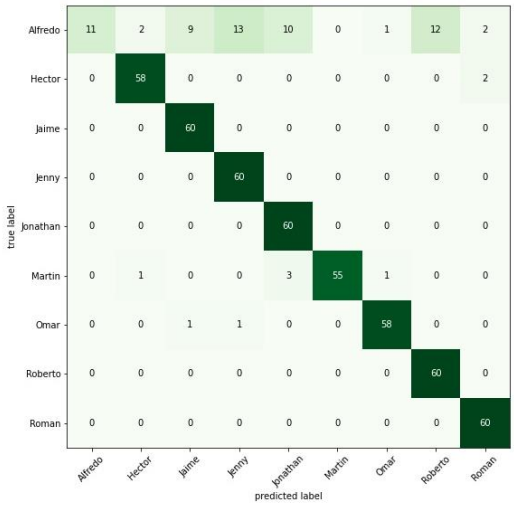


Figura 55, Evaluación matriz de confusión para RESTNET50

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

PRIMERO:

La detección facial y el reconocimiento facial funcionan juntos, pero no son la misma acción. El reconocimiento comienza con la detección, la detección es esencialmente el primer paso para el reconocimiento, distinguiendo las caras humanas de otros objetos en una imagen.

SEGUNDO:

Se logró realizar las pruebas de entrenamiento de forma correcta, comparando y dando mejores resultados los activadores de sofmax para los modelos utilizados VGG16, VGG19 y RESTNET50.

TERCERO:

Se demuestra que en el proceso de entrenamiento el modelo RESTNET50, es más rápido que los modelos VGG16 y VGG19, como se demuestra en la tabla 16 de resultados, sin embargo, no es más preciso como el VGG16 ni como el VGG19, existiendo una variación de 0.00102 del VGG19 frente al 0.2240 de RESTNET50 y de 0.00014 en el caso de VGG16 con la variante.

TERCERO:

Se demuestra que en el modelo VGG16 es más preciso y con el mínimo de error frente a los modelos VGG19 y RESTNET50, cumpliendo así la bibliografía presentada en el marco teórico "Very Deep Convolutional Networks for Large-Scale Image Recognition" [2], decidiendo para una implementación real el modelo **VGG16 confirmando su precisión mayor al 98%**.

CUARTO:

Haar Cascade Detection es uno de los algoritmos de detección de rostros más antiguos pero potentes inventados. Las características de Haar no solo se utilizan para detectar rostros, sino también para ojos, labios, matrículas de autos.

QUINTO:

La propuesta de solución garantizará la autenticidad, integridad, confidencialidad (información sensible) y no repudiación en el intercambio de información de trámites y servicios, utilizando identificación biométricos a través de X-ROAD.

SEXTO:

Se demuestra por la matriz de confusión, que la evaluación de desempeño de los modelos VGG16 y VGG19, son los más óptimos para una implementación en producción.

RECOMENDACIONES

Como trabajo futuro, se puede mejorar el rendimiento y los tiempos utilizando GPU's para el entrenamiento de las redes neuronales y dar la posibilidad de incluir más imágenes y más clases.

Es importante incluir para la parte de los micro-servicios realizar los despliegues en cluster de contenedores como los Kubernetes, esto ayudara a soportar la escalabilidad global del estado peruano.

Se recomienda utilizar mecanismos de presencia de vida, evaluando las micro-expresiones como se describen en [21], así como la utilización de otros medios biométricos como el reconocimiento de voz y reconociendo por iris.

Es importante considerar los temas éticos y los principios de privacidad de información biométrica, no se puede recolectar la información biométrica mientras no exista normativas y políticas públicas para escenarios de identidad biométrica.

REFERENCIAS BIBLIOGRÁFICAS

[1] Nam Jun Pyun. Extraction of an image in order to apply face recognition methods. Artificial Intelligence [cs.AI]. Université Sorbonne Paris Cité, 2015. English. ffNNT : 2015USPCB132, link encontrado en: https://tel.archives-ouvertes.fr/tel-01578110/document

[2] Karen Simonyan, Andrew Zisserman (2015), Very Deep Convolutional Networks for Large-Scale Image Recognition (ICLR 2015), link encontrado en: https://arxiv.org/abs/1409.1556

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun,Deep (2015) Residual Learning for Image Recognition, link encontrado en: https://arxiv.org/pdf/1512.03385.pdf

[4]. Gary Bradski (2015), Willow Garage, OpenCV, ROS, And Object Recognition ICRA Semantic Perception Workshop, link encontrado en: https://docs.opencv.org/5.x/da/d60/tutorial_face_main.html, http://www.ais.uni-bonn.de/~holz/spme/talks/01_Bradski_SemanticPerception_2011.pdf

[5] Marc Novakouski, Grace A. Lewis (January 2012), Interoperability in the e-Government

Context, TECHNICAL NOTE CMU/SEI-2011-TN-014. Carnegie Mellon, link encontrado en https://resources.sei.cmu.edu/asset_files/TechnicalNote/2012_004_001_15365.pdf

[6] Martin Kasparick, Stefan Schlichting, Frank Golatowski1, and Dirk Timmermann (2015), New IEEE 11073 Standards for interoperable, networked Point-of-Care Medical Devices.

[7]. Paul Viola, Michael Jones(2001), Rapid Object Detection using a Boosted Cascade of Simple Features, link encontrado en https://web.iitd.ac.in/~sumeet/viola-cvpr-01.pdf

[8]. Srivastava, Nitish (2014), ”Dropout: a simple way to prevent neural networks from overfitting”, JMLR 2014Srivastava, Nitish, et al. ”Dropout: a simple way to prevent neural networks from overfitting”, JMLR.

[9]. Ian Goodfellow, Yoshua Bengio, Aaron Courville, (2016) Deep Learning (Adaptative Computation and Machine Learning), MIT, link encontrado en: https://www.deeplearningbook.org/

[10]. Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow,

Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner,Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng (2015). TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. Link encontrado en: https://www.tensorflow.org/extras/tensorflow-whitepaper2015.pdf.

[11]. [Chyp Huyen, (January 2017) TensorFlow for Deep Learning Research, CS 20S, Stanford, link encontrado en: https://github.com/HaochunLiang/Stanford-CS20

[12]. Avinash Lakshman, Prashant Malik (2009), Facebook - Cassandra, A Decentralized Structured Storage System. link encontrado en: https://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf

[13]. Qi Cao, Ruishan Liu (2015), Department of EE Stanford University, Real-Time Face Tracking and Replacement, link encontrado en: https://web.stanford.edu/class/cs231m/projects/final-report-cao-liu.pdf

[14] Identity Authentication through Face Recognition (2019), link encontrado en: https://www.facefirst.com/

[15] Zuheng Ming, Junshi Xia, Muhammad Muzzamil Luqman, Jean-Christophe Burie, Kaixing Zhao (2019), FaceLiveNet+: A Holistic Networks For Face Authentication Based On DynamicMulti-task Convolutional Neural Networks. Cornell University, link encontrado en: https://arxiv.org/pdf/1902.11179.pdf

[16]. Giuseppe Garofalo, Vera Rimmer, Tim Van hamme, Davy Preuveneers, and Wouter Joosen, imec-DistriNet, KU Leuven (2018), Fishy Faces: Crafting Adversarial Images to Poison Face Authentication, link encontrado en:

https://www.usenix.org/system/files/conference/woot18/woot18-paper-garofalo.pdf

[17] Upal Mahbub, Vishal M. Patel, Deepak Chandra, Brandon Barbello, Rama Chellappa(2016), Partial Face Detection for Continuous Authentication, Cornell University, link encontrado en: https://arxiv.org/pdf/1603.09364.pdf

[18] Arman Boehm, Dongqu Chen, Mario Frank, Ling Huang,Cynthia Kuo, Tihomir Lolic, Ivan Martinovic,Dawn Song (2013),SAFE: Secure Authentication with Face and Eyes, University of California, Berkeley;Intel Labs;Nokia Research;Oxford University;Yale University, link encontrado en: https://people.eecs.berkeley.edu/~dawnsong/papers/safeI3e.pdf

[19] Xiaogang Wang, Yi Sun, Ziwei Liu, Zhenyao Zhu, Ping Luo, Xiaoou Tang (2014), DeepID:Deep Learning for FaceRecognition, Department of Electronic Engineering, The Chinese University of Hong Kong, link encontrado en: http://www.ee.cuhk.edu.hk/~xgwang/DeepID.pdf

[20] European Union Agency for Fundamental RightsHelping to make fundamental rights a reality for everyone in the European Union (Nov, 2019), facial-recognition-technology-fundamental-rights-considerations-law-enforcement, link encontrado en:

https://fra.europa.eu/en/news/2019/facial-recognition-technology-fundamental-rights-considerations-law-enforcement

[21] S. D. Lalitha,K. K. Thyagharajan (2019), Micro-Facial Expression Recognition Based on Deep-Rooted Learning Algorithm, International Journal of Computational Intelligence Systems, link encontrado en: https://doi.org/10.2991/ijcis.d.190801.001

[22] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger (2018), Densely Connected Convolutional Networks, link encontrado en: https://arxiv.org/pdf/1608.06993.pdf

[23] Kunihiko Fukushima (1980), Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybernetics 36, 193–202 (1980). link encontrado en: https://www.rctn.org/bruno/public/papers/Fukushima1980.pdf

[24] Asifullah Khan, Anabia Sohail, Amna Ali (2018), A New Channel Boosted Convolutional Neural Network using Transfer Learning, link encontrado en:

https://arxiv.org/ftp/arxiv/papers/1804/1804.08528.pdf

[25] Mei Wang, Weihong Deng (2020), Deep Face Recognition: A Survey, link encontrado en: https://arxiv.org/pdf/1804.06655.pdf

[26] Alex Polacco, Kayla Backes (2017), "THE AMAZON GO CONCEPT: IMPLICATIONS, APPLICATIONS, AND SUSTAINABILITY", Amazon Go Grocery(2020), Amazon Go(2016), link encontrado en:

https://www.amazon.com/b?ie=UTF8&node=16008589011,
http://wdsinet.org/Annual_Meetings/2017_Proceedings/CR%20PDF/cr215.pdf,

https://www.fujitsu.com/us/Images/Market-Insight_Amazon-Go.pdf

[27] Ryan Gross, (2016), A deep dive on how Amazon used AWS, ML, RL, and Simulation to power the Go Store's "Just Walk Out" experience, link encontrado en: https://towardsdatascience.com/how-the-amazon-go-store-works-a-deep-dive-3fde9d9939e9

[28] Lex Fridman, Research Scientist(2018), Computer Vision (2018), link encontrado en: https://www.dropbox.com/s/a5iqp1df5j1xydj/computer_vision_2018.pdf?dl=0

[29] James Cowley (2018), Redes neuronales convolucionales, IBM, link encontrado en: https://www.ibm.com/developerworks/ssa/library/cc-convolutional-neural-network-vision-recognition/index.html

[30] Shang-Hung Lin (2000), IC Media Corporation An Introduction to Face Recognition Technolog, link encontrado en: https://core.ac.uk/download/pdf/26858086.pdf

[31] TAHMID HASAN FUAD, AWAL AHMED FIME,DELOWAR SIKDER, MD. AKIL RAIHAN IFTEE, JAKARIA RABBI, MABROOK S. AL-RAKHAMI (MEMBER, IEEE), ABDU GUMAE, OVISHAKE SEN, MOHTASIM FUAD, MD. NAZRUL ISLAM (2021), Recent Advances in Deep Learning Techniques for Face Recognition, link encontrado en: Recent Advances in Deep Learning Techniques for Face Recognition (researchgate.net)

[32] Chip Huyen (2017), CS 20SI: Tensorflow for Deep Learning Research (Marzo, 2017), link encontrado en:

https://web.stanford.edu/class/cs20si/2017/syllabus.html,
https://web.stanford.edu/class/cs20si/2017/index.html

[33] Yante Li, Jinsheng Wei, Yang Liu, Guoying Zhao(2021), Deep Learning for Micro-expression Recognition: A Survey, link encontrado en: https://arxiv.org/pdf/2107.02823.pdf

[34] Lex Fridman, Research Scientist(2020), Deep Learning: State of the Art (2020), link encontrado en:

https://lexfridman.com/files/slides/2020_01_06_deep_learning_state_of_the_art.pdf,
https://deeplearning.mit.edu/

[35] Eugene Charniak (2018), Introduction to Deep Learning (2018), The MIT Press, Cambridge, Massachusetts, London, England., link encontrado en:

https://cs.brown.edu/courses/csci1460/assets/files/deep-learning.pdf

[36]TESLA, Piloto automático (2018), link encontrado en:

https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/solutions/resources/documents1/nvidia-m60-datasheet.pdf

[37] Ankith Jain Rakesh Kumar, Rajkumar Theagarajan, Omar Peraza and Bir Bhanu (2019), CLASSIFICATION OF FACIAL MICRO-EXPRESSIONS USING MOTION MAGNIFIED

EMOTION AVATAR IMAGES, Center for Research in Intelligent Systems, University of California, link encontrado en:

http://openaccess.thecvf.com/content_CVPRW_2019/papers/Face%20and%20Gesture%20Analysis%20for%20Health%20Informatics/Kumar_CLASSIFICATION_OF_FACIAL_MICRO-EXPRESSIONS_USING_MOTION_MAGNIFIED_EMOTION_AVATAR_IMAGES_CVPRW_2019_paper.pdf

[38] Mohammad Sabik Irbaz, MD Abdullah Al Nasim, Refat E Ferdous(2021),Real-Time Face Recognition System for Remote Employee Tracking, link encontrado en:

https://arxiv.org/pdf/2107.07576.pdf

[39] Ben Wycliff Mugalu, Rodrick Calvin Wamala, Jonathan Serugunda, Andrew Katumba (2021), Face Recognition as a Method of Authentication in a Web-Based System, link encontrado en: https://arxiv.org/pdf/2103.15144.pdf

[40] Z. Ming, J. Chazalon, M. Muzzamil Luqman, M. Visani and J. Burie (2018), "FaceLiveNet: End-to-End Networks Combining Face Verification with Interactive Facial Expression-Based Liveness Detection," 2018 24th International Conference on Pattern Recognition (ICPR), 2018, pp. 3507-3512, doi: 10.1109/ICPR.2018.8545274. link encontrado en: https://www.researchgate.net/publication/329312776_FaceLiveNet_End-to-End_Networks_Combining_Face_Verification_with_Interactive_Facial_Expression-Based_Liveness_Detection

[41] Evani Radiya-Dixit , Florian Tramèr (2021), Data Poisoning Won't Save You From Facial Recognition, link encontrado en: https://arxiv.org/pdf/2106.14851.pdf

[42] Delina Beh Mei Yin, Shariman Omar, Bazilah A. Talip (2017), Fusion of Face Recognition and Facial Expression Detection for Authentication: A Proposed Model, Malaysian Institute of Information Technology, link encontrado en: https://www.semanticscholar.org/paper/Fusion-of-face-recognition-and-facial-expression-a-Yin-Omar/87ee56feefdb39938cda7f872e784d9d986713af

[43] Yann Rodriguez and Sebastien Marce (2006), Face Authentication Using Adapted LocalBinary Pattern Histograms, IDIAP Research Institute, link encontrado en:

https://publications.idiap.ch/downloads/papers/2006/rodrig-eccv-2006.pdf

[44] Jerome Letier (2020), Alicem, Authentification en Ligne Certifiée sur Mobile (2020), link encontrado en: https://www.dw.com/en/france-embraces-facial-recognition-tech/a-51106489

[45] Apple, Face ID, autenticación facial en dispositivos móviles, link encontrado en: https://support.apple.com/en-us/HT208108, https://machinelearning.apple.com/research/face-detection

[46] Sermanet, Pierre, David Eigen, Xiang Zhang, Michael Mathieu, Rob Fergus y Yann LeCun (2013). OverFeat: reconocimiento, localización y detección integrados mediante redes

convolucionales. *arXiv: 1312.6229* [Cs], diciembre de 2013, link encontrado en: https://arxiv.org/abs/1312.6229

[47] Pawan Sinha, Benjamin Balas, Yuri Ostrovsky, and Richard Russell(2006), Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About, link encontrado en: http://web.mit.edu/sinhalab/Papers/19results_sinha_etal.pdf

[48] Ethan Meyers (2005) MIT, Face Databases, link encontrado en:

http://web.mit.edu/emeyers/www/face_databases.html, http://www.cs.cmu.edu/~cil/v-images.html

[49] Pete Fussey & Dr. Daragh Murray (2019),Independent report on the London Metropolitan Police Services Trial of Live Facial Recognition Technology, link encontrado en: http://repository.essex.ac.uk/24946/1/London-Met-Police-Trial-of-Facial-Recognition-Tech-Report-2.pdf

[50] Electronic Privacy Information Center - epic.org (2021), State Facial Recognition Policy, link encontrado en: https://epic.org/state-policy/facialrecognition/

[51] Nissenbaum, Helen F. and Introna, Lucas D.(2009), Facial Recognition Technology: A Survey of Policy and Implementation Issues (July, 22 2009). Center for Catastrophe Preparedness and Response, New York University, link encontrado en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1437730

[52] European Union Agency for Fundamental Rights (2020), Facial recognition technology: fundamental rights considerations in the context of law enforcement: link encontrado en: https://fra.europa.eu/en/publication/2019/facial-recognition-technology-fundamental-rights-considerations-context-law

[53] BBC News Mundo (2019), La polémica en China por la imposición del reconocimiento facial a todos los compradores de teléfonos, link encontrado en: https://www.bbc.com/mundo/noticias-50622301

[54] Philipp Terhorst, Daniel Fahrmann,Naser Damer,Florian Kirchbuchner and Arjan Kuijper(2020), Beyond Identity: What Information Is Stored in Biometric Face Templates?(2020), link encontrado en: https://arxiv.org/pdf/2009.09918.pdf

[55] INE (Instituto Nacional Electoral), ACUERDO DEL CONSEJO GENERAL DEL INSTITUTO NACIONAL ELECTORAL POR EL QUE SE APRUEBAN LAS ADECUACIONES PARA AMPLIAR Y FORTALECER EL SERVICIO DE VERIFICACIÓN DE DATOS DE LA CREDENCIAL PARA VOTAR (2020), link encontrado en: https://repositoriodocumental.ine.mx/xmlui/bitstream/handle/123456789/113983/CGex202005-15-ap-2-Gaceta.pdf, https://www.gob.mx/tramites, https://www.gob.mx/efirma, https://centralelectoral.ine.mx/2020/01/22/tus-datos-biometricos-debidamente-resguardados-ine/, https://www.gob.mx/conamer, https://www.senado.gob.mx/64/gaceta_del_senado/documento/28987

[56] Florian Schroff, Dmitry Kalenichenko, James Philbin (2015), FaceNet: A Unified Embedding for Face Recognition and Clustering, link encontrado en: https://arxiv.org/pdf/1503.03832.pdf

[57] CS231n Convolutional Networks for Visual Recognition of Standford University. Link encontrado en: https://cs231n.github.io/convolutional-networks/

[58] Alexander Amini (January-2021), Introduction to Deep Learning,MIT 6.S191, Link encontrado en: http://introtodeeplearning.com/slides/6S191_MIT_DeepLearning_L1.pdf

[59] Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks", Machine Learning. Link encontrado en: http://image.diku.dk/imagecanon/material/cortes_vapnik95.pdf

[60] Roseth Benjamin, Reyes Angela, Farias Pedro, Porrúa Miguel,Villalba Harold, Acevedo Sebastián, Peña Norma, Estevez Elsa, Linares Lejarraga Sebastián, Fillottrani Pablo (2018), "El fin del trámite eterno: Ciudadanos, burocracia y gobierno digital" , Link encontrado en: https://publications.iadb.org/publications/spanish/document/El-fin-del-tr%C3%A1mite-eterno-Ciudadanos-burocracia-y-gobierno-digital.pdf

[61] Gene Itkis, Venkat Chandar, Benjamin Fuller, Joseph P. Campbell, and Robert K. Cunningham(2015),Iris Biometric Security Challenges and Possible Solutions, Link encontrado en: https://www.ll.mit.edu/sites/default/files/publication/doc/2018-04/2015_09_ItkisG_IEEE-SPM_IrisKey_FP.pdf

[62] Ramadan Gad, M. Zorkany, AYMAN EL-SAYED, Nawal El-Fishawy(2015),Multi-Biometric Systems: A State of the Art Survey and Research Directions, Link encontrado en: https://pdfs.semanticscholar.org/4534/953f7aa995649693b0d6f006715ffdee591c.pdf?_ga=2.86058200.1189645931.1637168881-1590751028.1635828950

[63] Government Office for Science (2018), Biometrics: a guide. The National Archives, Kew, London TW94DU, Link encontrado en:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/715925/biometrics_final.pdf

[64] Simon Ouellet,François Grondin, Francis Leconte, François Michaud (2014), Multimodal Biometric Identification System for Mobile Robots Combining Human Metrology to Face Recognition and Speaker Identification, Link encontrado en:

http://dx.doi.org/10.13140/2.1.2138.8320

[65] Afshine Amidi, Shervine Amidi (2018), CS229-Machine Learning, Link encontrado en: https://stanford.edu/~shervine/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks#classification-metrics

[66] William T. Freeman, Antonio Torralba (2017), Lecture 8, Learned feedforward visual processing Neural Networks, Deep learning, ConvNets, MIT CSAIL, 6.869: Advances in Computer Vision. Link encontrado en:

http://6.869.csail.mit.edu/fa17/lecture/lecture8CONVNETS.pdf

LINK DE INTERNET

[1] Haarcascades, https://github.com/opencv/opencv/tree/master/data/haarcascades

[2] VGG, https://www.robots.ox.ac.uk/~vgg/

[3] OpenCV, https://opencv.org/

[4] Tensorflow, https://www.tensorflow.org/

[5] Keras, https://keras.io/

[6] scikit-learn, https://scikit-learn.org/stable/

[7] Plataforma de Autenticación Nacional, https://idperu.reniec.gob.pe/site/

[8] Ministerio del interior de Francia, https://www.interieur.gouv.fr/actualites/actu-du-ministere/alicem-premiere-solution-didentite-numerique-regalienne-securisee

[9] Convolutional Neural Networks, https://srdas.github.io/DLBook/ConvNets.html, https://cs231n.github.io/convolutional-networks/

[10] Convolutional Neural Networks, https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks

[11] Confusion matrix, https://stanford.edu/~shervine/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks#classification-metrics,

[12] Confusion matrix scikit-learn,

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion_matrix.html, https://scikit-learn.org/stable/auto_examples/model_selection/plot_confusion_matrix.html

[13] SoftMax,
http://sintesis.ugto.mx/WintemplaWeb/01Neural%20Lab/03Classification/04SoftMax/index.htm, https://deepai.org/machine-learning-glossary-and-terms/softmax-layer, https://ocw.mit.edu/courses/aeronautics-and-astronautics/16-412j-cognitive-robotics-spring-2016/lecture-notes/MIT16_412JS16_L16.pdf

ANEXO A: Configuración del Modelo VGG16+3

Ejecutando los siguientes parámetros:

data: {'uuid': 'ee8d2d10-e0be-447d-a9b3-d2f597f37193', 'epochs': 5, 'model': 'VGG16'}

y llamando a la función de keras en python: **model.summary()**

CNN MODELO= VGG16

CANTIDAD DE CLASES: 9

Model: "functional_1"

_____

Layer (type) Output Shape Param #

==========

input_1 (InputLayer) [(None, 224, 224, 3)] 0

_____

block1_conv1 (Conv2D) (None, 224, 224, 64) 1792

_____

block1_conv2 (Conv2D) (None, 224, 224, 64) 36928

_____

block1_pool (MaxPooling2D) (None, 112, 112, 64) 0

_____

block2_conv1 (Conv2D) (None, 112, 112, 128) 73856

_____

block2_conv2 (Conv2D) (None, 112, 112, 128) 147584

_____

block2_pool (MaxPooling2D) (None, 56, 56, 128) 0

_____

block3_conv1 (Conv2D) (None, 56, 56, 256) 295168

_____

block3_conv2 (Conv2D) (None, 56, 56, 256) 590080

_____

block3_conv3 (Conv2D) (None, 56, 56, 256) 590080

_____

block3_pool (MaxPooling2D) (None, 28, 28, 256) 0

block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 512)	12845568
dense_1 (Dense)	(None, 512)	262656
dense_2 (Dense)	(None, 256)	131328
dense_3 (Dense)	(None, 9)	2313

==========

Total params: 27,956,553

Trainable params: 13,241,865

Non-trainable params: 14,714,688