**UNIVERSIDAD NACIONAL JOSE MARIA ARGUEDAS FACULTAD DE INGENIERÍA ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**



**Reconocimiento de Personas Requisitoriadas del Sistema de Recompensas.pe usando OpenCV y SVM**

**ASIGNATURA :** Reconocimiento de Patrones

**CICLO :** 2020-II

**DOCENTE :**  Ing. Carlos Castro Buleje

**Integrantes:**

* Rubén Denky Valles Martinez
* Elizabeth Flores Huamaní
* Mario Quispe Quispe

**FECHA: 10/02/2021**

# **ÍNDICE**

[**Introducción**](#_py7mv28xckai) **3**

[**Marco Teórico**](#_wgi2jsh6bwcl) **4**

[Persona requisitoriada](#_8jnib7m3v4ye) 4

[Seguridad ciudadana](#_vclg5tfjurpk) 4

[Recompensas.pe](#_ofscey4m7iic) 4

[Clasificadores](#_10j9d798q7xy) 5

[Redes Neuronales](#_tbk5esfdw5mg) 5

[Máquina de vectores de Soporte](#_9vpja58crukk) 5

[Regresión Logística](#_w63yq6h1jm38) 5

[K-Vecinos más Cercanos](#_yt9w477b5coc) 6

[Perceptrón Multicapa (MLP)](#_7nviklj5u26r) 6

[Python](#_fm67g7pgpvhj) 6

[**Objetivos**](#_5s6ctzhfa1s7) **6**

[Objetivos Generales](#_o41k9j3h545b) 6

[Objetivos específicos](#_sbchtza9wmbb) 6

[**Metodología**](#_5s6ctzhfa1s7) **7**

[Recolección y tratamiento de imágenes](#_lihum2sjjen5) 7

[Entrenamiento o procesamiento de imágenes](#_6c45sfkmispe) 7

[Desarrollo del sistema o interfaz](#_yuozd025t0lz) 8

[**Desarrollo**](#_5s6ctzhfa1s7) **8**

[Opencv:](#_4eq0ubx3gztz) 9

[Sklearn:](#_4r412mucfxnf) 9

[Scikit-image:](#_mq4wgea42xbc) 9

[Pandas:](#_dhllpke5nof2) 9

[Numpy:](#_nbs8h9fdqaad) 10

[Pyinstaller:](#_ppfp55t2ywbf) 10

[Joblib:](#_m4jc1tyzq3w1) 10

[**Resultados**](#_w10g3yi2t96u) **10**

[**Conclusiones**](#_5s6ctzhfa1s7) **11**

[**Trabajos Futuros**](#_5s6ctzhfa1s7) **12**

[**Bibliografía**](#_287lt35tcazv) **12**

# 

# 

# **Introducción**

La búsqueda de personas requisitoriadas en el Perú es un problema grande, muchos de los ciudadanos peruanos están cansados de la inseguridad ciudadana, desde que salen al exterior de sus viviendas. El problema es que existen muchos criminales sueltos, con una larga lista de antecedentes penales. Estos criminales siguen cometiendo crímenes en la sociedad como: corrupción, delincuencia, violaciones, etc. Para prevenir estos daños es mejor capturarlos a estos individuos, para que no sigan cometiendo crímenes. El objetivo principal de este trabajo es desarrollar un programa de reconocimiento de personas requisitoriadas del sistema de recompensas.pe.

El presente trabajo está estructurado en 9 apartados. En el apartado 1 “Introducción” se hace una presentación general, el problema abordado y el objetivo. En el apartado 2 “Antecedentes” se mencionan algunos trabajos de investigación similares desarrollados. En el apartado 3 “Marco teórico” se definen todos los conceptos relacionados a nuestro tema de estudio. En el apartado 4 ”Metodología” se explica de qué metodología se aplicó para la realización del trabajo. En el apartado 5 “Desarrollo” se explica paso a paso la construcción del software y el lugar donde lo aplicaremos. En el apartado 6 “Pruebas de desarrollo” se muestra pruebas realizadas al algoritmo y se explica el dataset utilizado. En el apartado 7 “Resultados” se muestra los resultados obtenidos con el modelo. En el apartado 8 “Conclusiones” se analizan si los resultados fueron buenos o malos según a las premisas evaluadas en el trabajo. En el apartado 9 “Trabajos futuros” se da algunas sugerencias para las posibles mejoras que se podría hacer en este trabajo . En el apartado 10 “Bibliografía” se muestra a los autores citados.

**Antecedentes**

**APLICACIÓN MÓVIL DE RECONOCIMIENTO FACIAL EN PERSONAS CON ANTECEDENTES DE ABUSO SEXUAL EN LA &5PROVINCIA DE ANDAHUAYLAS, APURÍMAC - 2018** **- Tesis para optar el título profesional: Cáceres (2018)**

En este trabajo se detectaron rostros de personas que cometieron abusos sexuales en la provincia de Andahuaylas, se utilizó opencv y Face API que ofrece diversas funcionalidades que se encuentran la detección de caras, verificación facial, identificación facial, búsqueda de caras similares y la agrupación facial. También se utilizó el método de extractor de características de una imagen el PCA que mediante el cual uno puede transformar cada imagen original de la formación que figura en una Eigenface correspondiente. Una característica importante de la PCA es que uno puede reconstruir cualquier imagen original desde el conjunto de entrenamiento mediante la combinación de los Eigenface.

**Reconocimiento facial mediante el uso de PCA y algoritmos de reconocimiento Viola-Jones – Artículo Científico: Becerra (2017).**

según el autor esta última herramienta se vuelve bastante útil ya que es capaz de dividir en clase las muestras de imágenes de entrenamiento. El algoritmo PCA reduce eficientemente la dimensión de las imágenes, lo que nos dará información más precisa para la toma de decisiones de la máquina. El PCA sigue tres pasos. El primer paso es crear una matriz de transformación usando las imágenes de entrenamiento. El segundo paso es disponer cada imagen dentro de los vectores de la matriz. Como último paso para que la imagen de prueba sea reconocida, esta se introduce en el subespacio de los vectores propios de las caras EigenFaces y se compara con el dominio de los vectores propios con las caras de entrenamiento.

En este trabajo como principal entrenamiento se utilizó Máquina de Vectores de Soporte teniendo en cuenta que anteriormente las imágenes fueron transformadas con el método PCA.

**API DE RECONOCIMIENTO DE MICROEXPRESIONES FACIALES UTILIZANDO ALGORITMOS DE FISHERFACE Y SUPPORT VECTOR MACHINE: Mardones (2017).**

Según el autor el reconocimiento facial y de microexpresiones se ha convertido en los últimos años en un área de investigación activa, que abarca diversas disciplinas tales como procesamiento de imágenes, reconocimiento de patrones, visión por computador y redes neuronales. Se utilizó algoritmo de Fisherface que se comporta mejor en imágenes de variación de iluminación como de pose. Se implementó una API de reconocimiento facial que utiliza el algoritmo Fisherface y el clasificador de SVM para la clasificación de microexpresiones oculares y además pueda detectar si se encuentra con cejas arriba o estado normal. Las pruebas realizadas utilizando el kernel lineal de SVM, dieron como resultado que el modelo de microexpresiones oculares posee una precisión del 92% con base de datos de 1,106 imágenes.

# **Marco Teórico**

Para comprender el problema es necesario conocer el aspecto teórico o las bases teóricas del cual se rige el problema y el desarrollo de la solución. Por tal motivo a continuación se define el significado de las palabras utilizadas para el desarrollo del trabajo.

## **Persona requisitoriada**

Según el diccionario de la Real academia Española (RAE), la palabra requisitoria significa “la decisión judicial en virtud de la cual se ordena la ubicación, aprehensión y conducción de grado o fuerza de una persona”, por lo tanto entendemos que una persona requisitoriada es aquella que está siendo buscada por el sistema de justicia.

## **Seguridad ciudadana**

(Wikipedia, 2021) un sitio web muy conocido a nivel mundial lo define como: “La acción integrada que desarrolla el Estado, con la colaboración de la ciudadanía y de otras organizaciones de interés público, destinada a asegurar su convivencia y desarrollo pacífico, la erradicación de la violencia, la utilización pacífica y ordenada de vías y de espacios públicos y, en general, evitar la comisión de delitos y faltas contra las personas y sus bienes”.

## **Recompensas.pe**

Es una página web peruana que tiene como objetivo mostrar a todas las personas requisitoriadas por el que se ofrece una recompensa monetaria a aquellos ciudadanos colaboradores que brinden información oportuna e idónea que permita la búsqueda, captura y/o entrega de miembros de una organización criminal, personas vinculadas a ella o que actúen por encargo de la misma, organizaciones terroristas, así como presuntos autores y partícipes de uno o más delitos.

## **Clasificadores**

Son un conjunto de algoritmos que tienen la función de clasificar datos según a las entradas, mayormente existen los clasificadores supervisados en el cual se enseña al algoritmo a reconocer cierto tipo de objeto, entre los más conocidos están la regresión Logística, árboles de decisión, máquinas de vectores de soporte, k-vecinos más cercanos y muchos más.

## **Redes Neuronales**

"Una red neuronal es un sistema compuesto de muchos elementos procesadores simples operando en paralelo, cuya función es determinada por la estructura de la red, fuerza en las conexiones y el procesamiento realizado por los elementos computacionales en los nodos." (DARPA, 1988)

## **Máquina de vectores de Soporte**

Es un algoritmo muy utilizado en el campo de inteligencia artificial, el cual tiene la capacidad de clasificar y predecir valores según al conjunto de datos dados. Usa hiperplanos o vectores para separar los distintos tipos de datos, se define los vectores de soporte como “Los puntos que definen el margen máximo de separación del hiperplano que separa las clases” (iartificial.net, 2020)

## **Regresión Logística**

Es una función o método matemático utilizado para problemas de clasificación. la página web (analiticslane.com) lo define como “es una técnica de aprendizaje automático que proviene del campo de la estadística”.

## **K-Vecinos más Cercanos**

K vecinos más cercanos es uno de los algoritmos de clasificación más básicos y esenciales en Machine Learning. Pertenece a los algoritmos de aprendizaje supervisada en problemas de clasificación.

## **Perceptrón Multicapa (MLP)**

Según al sitio web (Wikipedia) el perceptrón multicapa se define como: “es una red neuronal artificial (RNA) formada por múltiples capas, de tal manera que tiene capacidad para resolver problemas que no son linealmente separables, lo cual es la principal limitación del perceptrón (también llamado perceptrón simple).

## **Python**

Es un lenguaje de programación versátil multiplataforma y multiparadigma que se destaca por su código legible y limpio. Una de las razones de su éxito es que cuenta con una licencia de código abierto que permite su utilización en cualquier escenario.

# **Objetivos**

## **Objetivos Generales**

Desarrollar un software de reconocimiento facial que sea capaz de reconocer a las personas requisitoriadas a nivel del departamento de Apurimac.

## **Objetivos específicos**

* Diseñar un software que sea capaz de procesar imágenes de diferente tamaño.
* Implementar el software que sea capaz de procesar imágenes en busca de personas requisitoriadas.
* Implementar el software que sea capaz de procesar video (webcam en tiempo real) en busca de personas requisitoriadas.
* Realizar pruebas tanto con imágenes en foto o a través de cámara.

# **Metodología**

Para un mejor entendimiento de este presente trabajo, se estableció una metodología conformada por tres etapas.

## **Recolección y tratamiento de imágenes**

Para la recolección de imágenes se tomó cien fotos personales a todos los integrantes de este proyecto como ejemplo para entrenar nuestro algoritmo.

En este proyecto se incorporó imágenes como ruido para que nuestro algoritmo reconozca mejor, que diferencie de la persona inocente y de la persona requisitoriada.

## **Entrenamiento o procesamiento de imágenes**

Para leer una imagen y posteriormente procesarla se está utilizando la librería OpenCV (Open Source Computer Vision) que es una librería open source de visión por computador que sirve para análisis de imagen.

Las imágenes obtenidas para el entrenamiento se está considerando de un tamaño de 128x128 que es conveniente para extraer mayores características de caras de las personas.

Luego se convierten las imágenes de RGB a escala de gris para posteriormente serializarlo cada imagen en un vector.

Antes de poder entrenar las imágenes se separó el 67% del total para el entrenamiento y el 33% del total para la prueba.

Para el entrenamiento se utilizó el algoritmo Máquina de Vectores de Soporte (SVM) debido a que se comporta mejor cuando hay varias imágenes en una foto a comparación con otros algoritmos.

El score obtenido con el algoritmo SVM es 0.981 que considerando los ruidos nos da un resultado muy alto.

## **Desarrollo del sistema o interfaz**

Para la parte del diseño del interfaz se utilizó el software QT Designer la versión 5.

Para la generación de la app de este proyecto se utilizó una librería llamada **Pyinstaller** que básicamente es una librería de Python para generar ejecutables .EXE que esta está dentro de la carpeta **dist**.

# **Desarrollo**

Para el desarrollo de software como lenguaje de Programación se toma Python ya que es muy utilizado en el desarrollo de algoritmos de Machine Learning, motivo por el cual existen diferentes librerías, herramientas y una documentación amplia que nos pueden ayudar a conseguir los objetivos de una manera más sencilla, la aplicación tendrá 5 fases principales:



Para la recolección y el procesamiento de imágenes se utilizó la librería de Opencv-python.

Para el modelo entrando de el dataset fue mediante el uso de el algoritmo de de Machine Learning, Máquina de Vectores de Soporte (SVM), utilizando la librería La librería Scikit Learn que proporciona un modelo bastante completo para implementar el algoritmo de Máquinas Vectores de Soporte Clasificación y para definir el módulo será: sklearn.svm.

Para elegir el uso del algoritmo más indicado para la resolución de este problema, se realizó una comparación de los siguientes algoritmos: SVM, KNN, MLP, RL. y como mejor algoritmo con mayor exactitud de mejores resultados es el SVM.

A continuación se mencionan los elementos que se tuvieron en cuenta en cada aspecto y del desarrollo y se detallan los mismos.

Las herramientas a utilizar para el desarrollo del software se mencionan a continuación:

## **Opencv:**

Es una librería open source de visión por computador, análisis de imagen y aprendizaje automático.

## **Sklearn:**

Es una biblioteca de código para el análisis de datos contiene una serie de implementaciones para diferentes algoritmos populares de aprendizaje automático.

## **Scikit-image:**

Es una colección de algoritmos para el procesamiento de imágenes, para filtrar una imagen.

## **Pandas:**

Es una librería de python destinada al análisis de datos, que proporciona unas estructuras de datos flexibles y que permiten trabajar con ellos de forma muy eficiente.

## **Numpy:**

Es una biblioteca que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.

## **Pyinstaller:**

Es una biblioteca que servirá para generar el punto ejecutable de nuestra aplicación.

## **Joblib:**

Es una biblioteca que servirá para generar y cargar archivos de modelo entrenado.

El presente informe tiene como objetivo principal desarrollar una aplicación de Reconocimiento de personas requisitoriadas del sistema de recompensa.pe usando SVM.

Esta aplicación será implementada en la región de Apurímac. La operatividad de la aplicación consiste en que el usuario normal realice una captura de la persona requisitoriada en tiempo real con la cámara encendida, a la vez tiene la opción de cargar la foto en donde el requisitoriado se encuentre entre multitud de personas.

Una vez procesada la imagen el software identifica el rostro de la persona buscada, mostrando un mensaje si es requisitoriado o inocente, en cuestión de segundos.

# **Resultados**

En la tabla 01 se muestra los resultados del accuracy de los distintos algoritmos con respecto a las diferentes clases dadas. Se puede notar que el algoritmo que tuvo el mayor porcentaje de aciertos fue el de Regresión Logística. Para este proyecto, se utilizó el algoritmo de máquinas de vectores de soporte debido a que al momento de probar la aplicación haciendo el análisis en tiempo real, se notó que tenía mejores resultados que la regresión logística. Además, al cargar con datos que tenían más de dos personas, el algoritmo de svm también tuvo mejores resultados que la regresión logística.

Tomando A= cantidad de aciertos y F = cantidad de fallos, se muestra la siguiente tabla:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Denky** | | **Elizabeth** | | **Mario** | | **Ruido** | | **Porcentaje de aciertos** |
|  | **A** | **F** | **A** | **F** | **A** | **F** | **A** | **F** |
| **SVM** | 31 | 3 | 30 | 0 | 34 | 0 | 67 | 0 | 98.1% |
| **KNN** | 34 | 0 | 30 | 0 | 34 | 0 | 53 | 14 | 91.5% |
| **MLP** | 29 | 5 | 30 | 0 | 34 | 0 | 63 | 4 | 94.5% |
| **RL** | 34 | 0 | 30 | 0 | 34 | 0 | 65 | 2 | **98.7%** |

**Tabla 01: Comparativa del resultado de los distintos algoritmos de clasificación**

# **Conclusiones**

* Como conclusión principal se desarrolló un software de reconocimiento facial que sea capaz de reconocer a las personas requisitoriadas a nivel del departamento de Apurimac.
* Se desarrolló una interfaz amigable para que cualquier usuario común, pueda entenderlo de cómo debe usar.
* Como mejor algoritmo que da mejores resultados en el accuracy (exactitud) y en el testing es el SVM.
* Se diseñó un software que sea capaz de procesar imágenes de diferentes tamaños.
* Se implementó el software que sea capaz de procesar imágenes en tiempo real.
* Se desarrolló un análisis multiusuario, es decir se identificó a la persona requisitoriada entre multitud de personas.

# **Trabajos Futuros**

* En futuras versiones, se podría implementar la opción de trabajar con un conjunto grande de imágenes, el cual facilitará la búsqueda del requisitoriado.
* Para obtener una mayor precisión en el entrenamiento y prueba de los datos, sería conveniente usar algoritmos especializados en el análisis de imágenes los cuales son las redes neuronales convolucionales.
* La generación de una aplicación multiplataforma como android y ios es una buena opción ya que permitirá a que la población en general tenga acceso al análisis y por ende tener mayor posibilidad de ubicar a los requisitoriados.
* Las páginas web son una buena opción para llegar a la mayor cantidad de personas en la búsqueda, es implementar el sistema en una aplicación web en donde todos tengan acceso.
* Otra buena opción es implementar el sistema en cámaras de videovigilancia en las ciudades, haciendo un análisis en tiempo real de las imágenes.

# **Bibliografía**

Recuperado de <https://lpderecho.pe/que-es-requisitoria-como-saber-persona-tiene-requisitoria-exp-5060-2009-phd-tc/> el 09 de febrero de 2021.

Recuperado de <https://es.wikipedia.org/wiki/Seguridad_ciudadana> el 09 de febrero de 2021.

Recuperado de [https://recompensas.pe/](https://recompensas.pe/acerca-de) el 09 de febrero de 2021.

Recuperado de <https://www.usmp.edu.pe/publicaciones/boletin/fia/info32/pag4.htm> el 09 de febrero de 2021.

Recuperado de <https://www.iartificial.net/maquinas-de-vectores-de-soporte-svm/> el 09 de febrero de 2021.

Recuperado de <https://www.analyticslane.com/2018/07/23/la-regresion-logistica/> el 09 de febrero de 2021.

Recuperado de <http://181.176.178.114/handle/123456789/358> el 09 de febrero de 2021

Recuperado de <https://www.academia.edu/download/57330261/titulo-magico-sobre.pdf> el 09 de febrero de 2021

Recuperado de <http://dspace.utalca.cl/handle/1950/12266> el 09 de febrero de 2021

Recuperado de <https://scikit-image.org/> el 09 de febrero de 2021.

Recuperado de <https://robologs.net/tutoriales/tutoriales-opencv/> el 09 de febrero de 2021.