

Detección de individuos en acceso principal | Phytton & OpenCV



Colaboración

Uzziel Fuentes Enriquez, José Antonio Montero Valverde, Miriam Martínez Arroyo, José Francisco Gazga Portillo, Tecnológico Nacional de México / campus Acapulco

RESUMEN: La detección de individuos consiste en reconocer en imágenes obtenidas a través de un sensor (cámara) la presencia en tiempos diferentes del mismo individuo, esto mediante técnicas de visión artificial. A pesar de recientes avances en este campo, este problema aún permanece como un reto debido a que las imágenes obtenidas en algunas cámaras de video vigilancia (CCTV) suelen ser de baja calidad, presentan cambios en la iluminación, así como variaciones en las poses de las personas. Algunos métodos basados en aprendizaje profundo han alcanzado un notable avance en este tema, estos tienen como objetivo aprender las características. En la región de interés "RoI" en el acceso principal dada una imagen o secuencia de imágenes al ser tratadas y analizadas se puede lograr obtener aplicando con una metodología de visión artificial la detección en el acceso principal de la región de interés y de esta manera alertar oportunamente una posible situación de inseguridad.

PALABRAS CLAVE: CCTV, extracción de características, reconocimiento de patrones, RoI, seguridad, visión artificial.

ABSTRACT: The detection of individuals consists of recognizing in images obtained through a sensor (camera) the presence of the same individual at different times, this by means of artificial vision techniques. Despite recent advances in this field, this problem still remains a challenge due to the fact that the images obtained in some video surveillance cameras (CCTV) tend to be of low quality, present changes in lighting, as well as variations in the poses of people. Some methods based on deep learning have reached a remarkable advance in this subject, these aim to learn the characteristics. In the region of interest "RoI" in the main access given an image or sequence of images when treated and analyzed, it can be obtained by applying with an artificial vision methodology the detection in the main access of the region of interest and in this way promptly alert a possible situation of insecurity.

KEYWORDS: CCTV, feature extraction, pattern recognition, RoI, security, machine vision.

INTRODUCCIÓN

En los últimos años, la necesidad de mejorar la seguridad en diferentes contextos de la sociedad actual, dan lugar al desarrollo de nuevo hardware y software informático empleando técnicas de detección automática basadas en rasgos biométricos. [1]

El tema de seguridad aplicando visión artificial es de gran interés para diferentes áreas de comunidades sociales y evidentemente la detección de individuos es fundamental, porque apoya a prevenir y combatir eventos como secuestros, asaltos, hurtos, incluso actividades clasificadas como de terrorismo que han mostrado la necesidad de desarrollar un método confiable para detectar a un individuo en tiempos diferentes y así evitar acciones que compro-

meten la seguridad, en este artículo se muestra mediante técnicas de visión artificial, algunos resultados alcanzados al detectar la presencia de un individuo en un acceso principal. [2]

Cuando se intenta llevar a cabo control de seguridad en un acceso principal mediante un sistema de visión artificial se plantea un inconveniente muy serio, la cantidad de información que se debe analizar para poder realizar la detección. Pues llevar a cabo acciones que para un humano pueden resultar excesivamente largas manualmente, mediante un sistema de visión artificial de toma de decisión y predicción automático se puede simplificar. [3]

Dadas las técnicas que existen en la actualidad de visión artificial, es posible pensar que se puedan usar algunas de ellas para ayudar con el desarrollo de un sistema. Los métodos de inteligencia artificial que tienen una mayor utilización son los sistemas expertos, a través de los cuales se pueden seguir las acciones que se deben realizar. Son importantes los métodos de clasificación, sobre todo para elegir de la información que se genera sólo aquella que sea adecuada para el análisis de la seguridad, entre los métodos de clasificación destacan los de redes neuronales ver figura 1. [4]

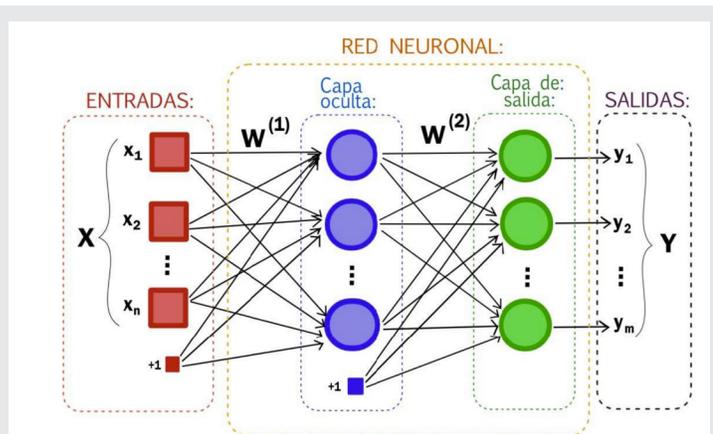


Figura 1. Red Neuronal, grafo dirigido.
Fuente:Z

Mientras tanto los sistemas biométricos son sistemas automatizados que permiten la identificación y verificación de individuos de forma rápida y segura, puesto que solo se analizan determinados patrones biométricos que no pueden ser alterados, manipulados, falsificados o robados para recrear información personal o acceder a la información de otros. [5]

Estos patrones son características tanto físicas como de comportamiento. Las huellas dactilares, la retina, el iris, los patrones faciales, de venas de la mano o la geometría de la palma de la mano, representan ejemplos de características físicas, mientras que en las características del comportamiento se incluye el reconocimiento de voz, firma, dinámica del teclado, cadencia del paso (paseo) y conducta gestual.

Todos estos son rasgos inherentes al individuo, es por ello que los sistemas inteligentes suponen una ventaja frente a sistemas tradicionales que suelen requerir información adicional como documentos de identidad, tarjetas de acceso, que pueden ser extraviados o robados provocando brechas de seguridad.

Para una correcta detección se deben obtener algunas características llamados descriptores y son de distintos tipos, estos se pueden centrar en diferentes características de las imágenes como color, tamaño, textura, borde. Para ello con OpenCv y Python, utilizando técnicas de visión artificial como HoG, DoG, PCA, SVM, RandomForest, Violas&Jones, etc. [6]

MATERIAL Y MÉTODOS

Pese a que el problema en la detección de rostros es distinto a la detección de individuos, ambos tienen como objetivo detectar la ROI de la imagen obtenida por una cámara (sensor). En el primer caso, se tienen imágenes de rostros comúnmente en vista frontal con limitaciones como cambios de expresiones faciales, iluminación, maquillaje, envejecimiento, presencia de oclusiones tales como bufandas u objetos en frente, en el segundo caso (el de interés) se tienen muestras con variaciones en calidad dependiendo del clima, imágenes de cuerpo entero que posiblemente debido a la posición de la cámara el rostro podrá aparecer parcialmente o no ser enfocado. [7]

Objetivos.

- Implementar sensores de trabajo que corresponde a la parte física, ver figura 2.
- Definir la orientación del sensor, en el caso del acceso principal, un ángulo correcto de captura como se muestra en la figura 3, donde la posición de la cámara procura abarcar la zona de acceso principal para detectar la ROI, debe estar ubicada a 90° y con un movimiento horizontal frontal de $0 - 180^\circ$, aunque la posibilidad del equipo es de 360° y para el análisis solo interesa una sola orientación momentáneamente como se muestra en la figura 3 al horizonte en 90° no mayor de 45° descendiente horizontal.
- Realizar distintas pruebas con diversas técnicas de visión artificial y distintos entornos siguiendo la línea Python y OpenCv.
- Obtener imágenes, pre-procesarlas, obtener descriptores, segmentar, etc.
- Discriminar zonas que no sean de interés, de este modo solo interesara poder detectar al individuo.

Implementación del sensor de trabajo.

Para definir las características básicas de detección automática de la región de interés basado en las técnicas de visión artificial se deben seguir una serie de pasos importantes.

En el caso particular para la zona de acceso principal se implementó un sistema CCTV mediante cámaras PTZ instaladas en dos zonas estratégicas de la Unidad del Departamento de Investigación y Posgrado del Tecnológico

Nacional de México, Campus Acapulco (DEPI):

- Biblioteca
- Acceso principal (zona que interesa).

La implementación del CCTV mediante cámaras con sensores digitales Z-20A-2(IR) presenta las siguientes características técnicas que permiten la obtención de imágenes de estudio a realizar:

- E / S de alarma: 7 entradas / 2 salidas
- Soporte de CoC
- 4 en 1 (AHD / TVI / CVI / CVBS)
- Protocolo adaptativo PELCO-P / D.
- Recuperación de memoria.
- Control en tiempo real: llame a cualquier función en cualquier momento.
- Admite 255 posiciones preestablecidas.
- Menú OSD en inglés.
- Zoom proporcional.
- Función de volteo automático.
- Construido en protección contra rayos y sobretensiones.
- Motor pasó a paso avanzado para un movimiento más suave.
- Soporte Tour 4 grupos, 8 escaneo automático, 4 patrones y escaneo continuo de 360 ° y escaneo intermitente de 360°.
- El domo superior adopta una estructura de aleación de aluminio de 7 " para mejorar el tiempo de vida.



Figura 2. Ubicación de sensor en acceso principal.
Fuente: Elaboración propia del autor.

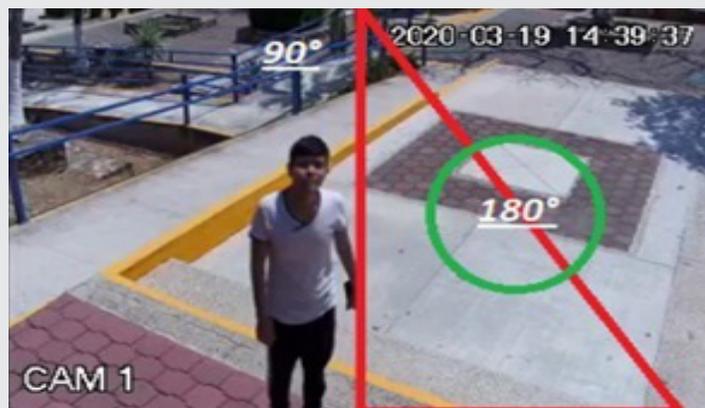


Figura 3. Orientación del sensor a la zona de interés.
Fuente: Elaboración propia del autor.

Definidas las características necesarias para trabajar, entre ellas, el dispositivo de grabación, como se muestra en la figura 4, utilizando para el funcionamiento correcto, un DVR no NVK, en este se configura la calidad de la imagen con la que se va a trabajar, esto es muy importante pues dependerá de esto el % de la tasa de éxito al eliminar posible ruido desde un inicio en la entrada de dato para el procesamiento de la detección automática.

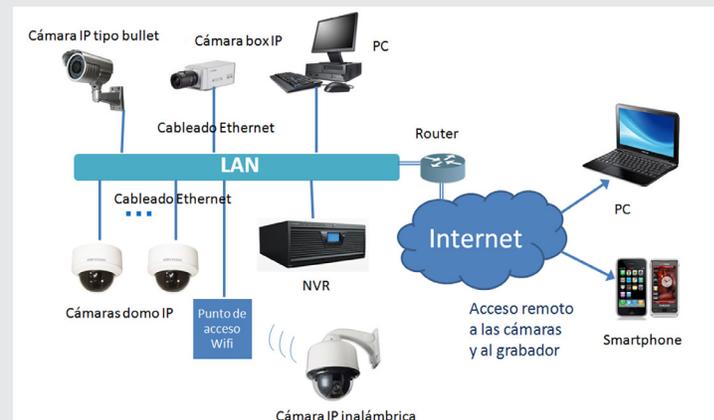


Figura 4. Diagrama genérico para DVR o NVK.
Fuente: Manual de usuario DVR dahua.

Una vez optimizado el medio de captura se extraen las imágenes ya sea en video o imagen estática, de lo cual dependerá el grado de complejidad del proceso de detección automática.

Metodología

Definida la imagen o imágenes en caso de un frame de video lo que procede es el tratamiento para mejorar posibles fallos en la calidad (esto no quiere decir que la imagen sea mala), uno de los principales pasos al iniciar a trabajar con una imagen $f(x, y)$ es trabajar con la vecindad de píxeles y para ello es conveniente reducir a escala de grises para agilizar el procesamiento (ver figura 5), esto no es realmente necesario si se cuenta con un equipo de cómputo adecuado, pero definitivamente disminuye el tiempo de proceso.

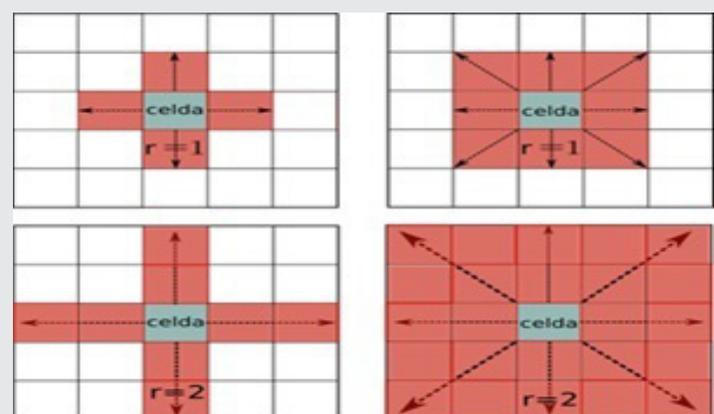


Figura 5. Vecindad para reducir imagen $f(x, y)$.
Fuente: <https://naukas.com/2020/08/03/inteligencia-artificial-y-automatas-celulares/>

Cuando la imagen esta tratada para ser procesada lo que prosigue es definir ciertas características que pueden delimitar la región de interés como son la distancia entre extremidades y para ello se puede usar un bordeado (contorno) y realizar un Sobel ver figura 6. Los operadores de Sobel se pueden ver como la combinación unidimensional de diferenciación y cierto suavizamiento.



Figura 6. Discriminación de la zona de interés.

Fuente: Elaboración propia del autor.

Las características “descriptores” más comunes para delimitar o segmentar regiones son: intensidad de los pixeles, textura, color y gradiente. Una suposición importante, que normalmente se asume en visión de nivel intermedio, es considerar que pixeles de un mismo objeto comparten propiedades similares. Por ejemplo, al procesar una imagen de un limón, suponemos que el color de sus pixeles es aproximadamente homogéneo (verdoso).

En la vida real esto no es totalmente cierto, el color de los pixeles varía. Para evitar este tipo de variaciones es mejor considerar un color “aproximadamente” similar sobre una región más que a nivel pixel. Esto no es un problema sencillo, ya que es difícil distinguir las variaciones propias del objeto o por cambios de iluminación (suponiendo, cuando hay sombras en los objetos, retomar la figura 6), de las diferencias por tratarse de otro objeto.

En el análisis de los objetos “individuos” en imágenes es esencial poder distinguir entre los objetos de interés y “el resto”, normalmente se hace referencia a este último grupo como el fondo. Las técnicas que se utilizan para encontrar los objetos de interés (RoI) se refieren generalmente como técnicas de segmentación (segmentar el primer plano del fondo). El resultado de esta segmentación da como resultado una imagen binaria. Normalmente se utiliza la convención de que se le asignan el valor 1 a los pixeles que corresponden a objetos y el valor de 0 a los pixeles que corresponden al fondo. Como resultado de dicha segmentación, la imagen es partida en regiones y se conocen los bordes entre las regiones. [6, p. 2]

Uno de los operadores de Sobel se obtiene como el producto de un vector de diferenciación (D) por uno de suavizamiento (G): Sobel = DG T.

Donde $D = (-1; 0; 1)$ y $G = (1; 2; 1)$. Esto reduce el efecto de aplicación del ruido que es característico de los operadores derivativos, por esto generalmente se prefiere el operador de Sobel.

Las técnicas clásicas de detección de orillas se basan en diferenciar a la imagen, esto es, encontrar la derivada respecto a los ejes (x, y) , o gradiente. El gradiente de una función $f(x, y)$ se define como:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \quad \text{Ec (1)}$$

La magnitud del gradiente (∂f) se calcula como:

$$|\nabla f| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad \text{Ec (2)}$$

De esta manera se puede discriminar la zona y lograr obtener los patrones de medida necesarios para definir el contorno o área del individuo. De modo que una vez obtenido este procesamiento ya sea mediante un análisis PCA o randomForest se logra reducir significativamente la imagen y así realizar una vez más un proceso en el cual ya se detectan las características “descriptores” que identifican mediante un clasificador el cual puede ser genérico o creado propiamente.

No esta de más mencionar que la reducción tomando el modelo básico RGB puede o no cambiar a escala de GRIS. [8]

Cada color se representa como un vector del origen y la diagonal principal corresponde a la escala de grises. En este modelo se basan las cámaras y receptores de televisión (TV). Sin embargo, se tienen problemas al aplicarlo a procesamiento de imágenes (ecualización) y visión (no-lineal).

Con todos estos datos se programa mediante OpenCv y Python la abstracción de la realidad y se realiza el análisis y detección de la región de interés. Mediante los patrones se puede clasificar, por eso es importante desde un inicio definir la región de interés a trabajar. Y de este modo se puede realizar una correcta segmentación. Que sirve para bordear la región y obtener la descripción.

El algoritmo recibe los datos de la imagen procesada, los identifica y decide lo que posiblemente puede ser. En algunos casos la sombra puede ser considerada como parte debido a que es una imagen simétrica reflejada, pero esto se corrige definiendo la de mayor tamaño (silueta).

Un sistema de visión artificial consta de forma general de un conjunto de subsistemas capaces de realizar las funciones de adquisición y análisis ver tabla.

Ahora bien, esto solo es la detección del individuo, la idea a futuro es detectarlo en tiempos diferentes [9], para ello se deben almacenar las imágenes y compararlas, si guardan similitudes es positiva y se re identifica al individuo en el acceso principal de manera “sospechosa” al aparecer en tiempos distintos en la zona.

Este proceso se describe de la siguiente manera, Entrada → Pre-procesamiento → Identificación → Base de datos → Re-identificación → Salida.

Describiendo los datos obtenidos mediante PCA se puede realizar una predicción usando la clasificación randomForest que es uno de los varios clasificadores, también con el uso de HOG. Finalmente decir que este proceso se puede representar de varias formas.

RESULTADOS

Desde el inicio al discriminar la zona de interés, de este modo solo interesara poder detectar a individuos mediante características, figura 7.

Para poder realizar esta fase y poder evidenciar se realizó un pequeño set de imágenes tomadas del área del acceso principal que es de donde se obtiene la región de interés, ver figura 8.

Una vez teniendo las imágenes se procede a programar, para el proceso de detectar regiones en movimiento simplemente se comparan dos imágenes y se debe encontrar las diferencias entre ambas (resta).

Con este principio básico y lógico este algoritmo hace algo similar utilizando una serie de métodos y técnicas que lo optimizan y simplifica, ya que comparar pixel a pixel no sería viable ni óptimo.

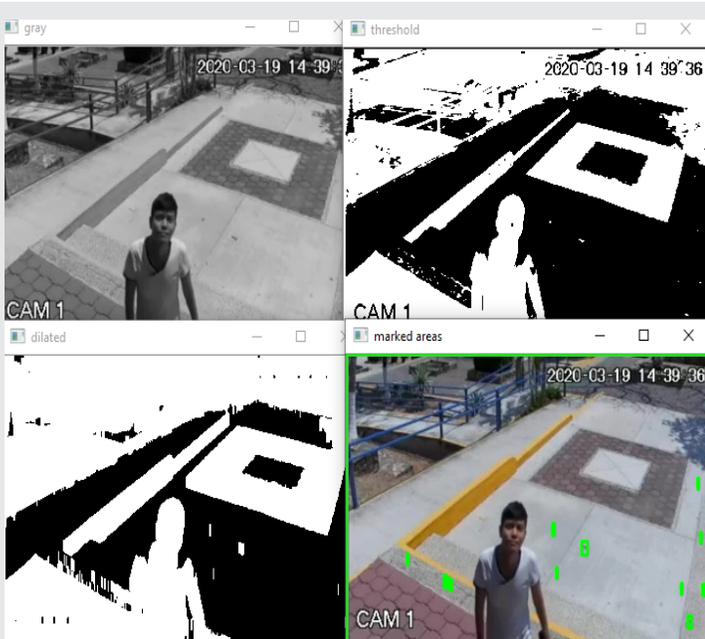


Figura 7 Discriminación de la zona para la región de interés. Fuente Elaboración propia del autor.



Figura 8. Set de 8 imágenes del acceso principal. Fuente: Elaboración propia del autor.

Tabla 1. Información de una de las imágenes redimensionada. Fuente: Elaboración propia en Python.

Image Dimension	: (240, 352, 3)
Image Height	: 240
Image Width	: 352
Number of Channels	: 3

Por lo tanto los pasos realizados serán:

- Capturar una secuencia de imágenes como mínimo, ocho.
- Se transforman las imágenes a una escala de grises (esto mejora el rendimiento del proceso, dependiendo del equipo de proceso es su uso) figura 10.
- Se reduce la resolución de las imágenes previas para mejorar el rendimiento a la hora de “comparar imágenes” (imágenes con resoluciones altas hace que el proceso sea más lento) por ello importante aplicar la reducción de la dimensionalidad, tabla 1.
- Se analizan las imágenes para extraer características figura 11.
- Se analiza la zona y se identifica la región de interés o se toman en cuenta aquellas que tienen una cierta relevancia, por ejemplo, si cambian 5 píxeles, pues igual no sea relevante y se identifica al individuo, figuras 12 y 13.

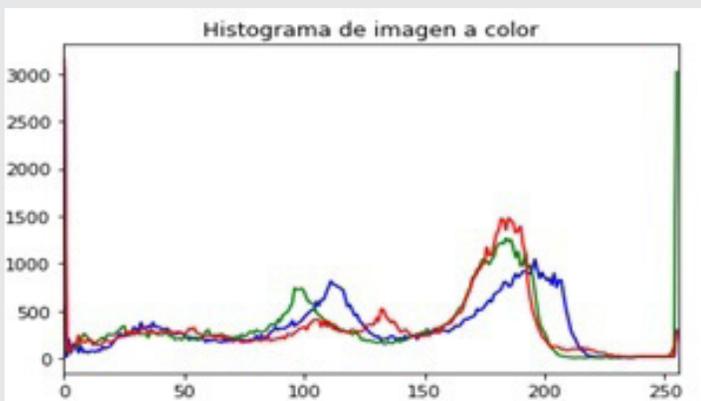


Figura 9. Histograma de canales RGB de la imagen analizada antes de transformarse a escala de gris. Fuente: Elaboración propia Python & OpenCv



Figura 12. Detección de individuos. Fuente: Elaboración propia.



Figura 10. Histograma en escala de grises intensidad lumínica y cantidad de pixeles. Fuente: Elaboración propia.

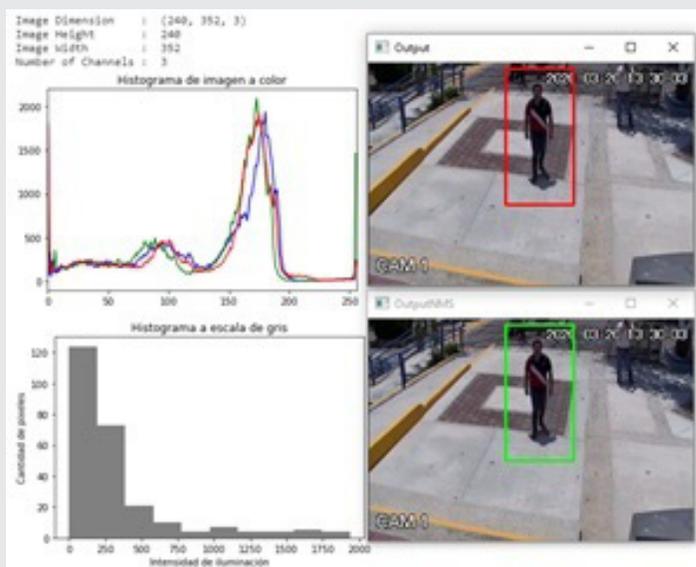


Figura 13. Detección de individuo. Fuente: Elaboración propia.

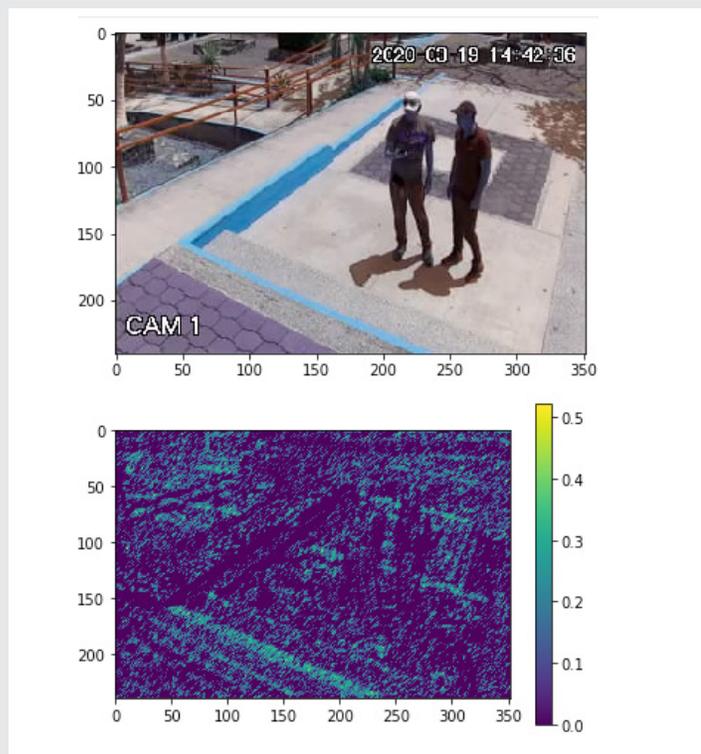


Figura 11. Aplicado HoG para conseguir características. Fuente: Creación propia del autor Python.

AGRADECIMIENTOS

Al apoyo brindado por CONACYT, el cual ha permitido implementar un prototipo en el Tecnológico Nacional de México Campus Acapulco, dentro del sistema de posgrado MSC DSI. Prototipo que para su fase final está contemplado para la STPS unidad regional Acapulco del estado de Guerrero.

CONCLUSIONES

Se concluye que PCA antes de RF (randomForest) no es una gran ventaja, si es que la hay. La regresión de componentes principales (PCR) es, por ejemplo, cuando, PCA ayuda a regularizar las características de entrenamiento antes de la regresión lineal y eso es muy necesario para conjuntos de datos dispersos. Como randomForest "RF" en sí ya realiza una regularización (buena / justa) sin asumir la linealidad, no es necesariamente una ventaja. Bajo tales circunstancias, es una pequeña ventaja pre filtrar con PCA, para poder lograr una tasa exitosa en la identificación.

BIBLIOGRAFÍA

- [1] J. Anil K, R. Arun A y N. Karthik, *Introducción to biometrics*, New Delhi, India: Springer US, 2012.
- [2] C. Gomez Cid, «<http://penta.ufrgs.br/>,» 24 08 2020. [En línea]. Available: <http://penta.ufrgs.br/>.
- [3] G. M. Marakas, *Decision support systems in the 21st century*, Upper Saddle River, NJ: Prentice Hall, ©2003.
- [4] A. Gonzáles Marcos, F. J. Martínez de Pisón Ascacibar, A. V. Pernía Espinoza, F. Alba Elías, M. Castejón Limas, J. Ordieres Meré y E. Vergara Gonzáles, *Técnicas y Algoritmos Básicos de Visión Artificial*, La Rioja, España.: Universidad de la Rioja España, ISBN 84-689-9345-X, 2015.
- [5] S. K. Modi, *Biometrics in Identity Management: Concepts to Applications*, Boston: Artech House ISBN: 1-60807-018-2, 2012.
- [6] A. Ramírez-Manzanares, *Reporte de Búsqueda, detección y conteo de objetos*, Apdo. Postal 402, Guanajuato, Gto., 36000, Mexico: Centro de Investigación en Matemáticas A.C, 2015.
- [7] D. Yi, Z. Lei, S. Liao y S. Z. Li, «Deep metric learning for person re-identification,» *International Conference on Pattern Recognition (ICPR)*, pp. 34-39, 2014.
- [8] M. A. Alonso Pérez y J. J. Báez Rojas, «Uso del sistema HSI para asignar falso color a objetos en imágenes digitales,» *REVISTA MEXICANA DE FÍSICA E 54 (2) 186-192*, pp. 186-192, 12 febrero 2008.
- [9] A. Ahmed Sekh, D. Prosad Dogra, H. Choi, S. Chae y I.-J. K. , «Person Re-identification in Videos by Analyzing,» *Multimedia Tools and Applications*, vol. 24537-24551, n° 79, p. 15, 23 Jun 2020.
- [10] «CERTSI, Blog, «La problemática de la biometría como medio de autenticación»,» [En línea]. Available: <https://www.certsi.es/blog/problematika-biometria-autenticacion>.
- [11] J. C. Moctezuma, *Neural Network toolbox de Matlab*, México: ciencias computacionales INAOE, septiembre 2006.
- [12] L. E. Sucar y G. Gomez, *Vision Computacional*, Puebla, Helmholtz Zentrum Munchen: Instituto Nacional de Astrofísica, Óptica y Electrónica.
- [13] «Cuerpo Nacional de Policía, Sistema Automático de Identificación Dactilar – SAID,» [En línea]. Available: http://www.policia.es/org_central/cientifica/servicios/id_identificacion.html. [Último acceso: 2 noviembre 2019].
- [14] H. T y R. S, «Object identification in a Bayesian context,» *International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1276-1282, 1997.
- [15] G. BATCHEN, *Arder en deseos.: La concepción de la fotografía*, Barcelona: Gustavo Gili, S.L., 2004.
- [16] L. C y . G. S., et al, *Person re-identi*, In Fusiello, A., Murino, V., et al.: *Computer Vision (ECCV)*, pages 391{401, 2012.