

Adaptación y comparación de dos metodologías de reconocimiento facial aplicadas a la detección de somnolencia en conductores

Johanna Gorvenia
jgorvenia@hotmail.com / Universidad de Lima. Lima, Perú

Fernando Tello
feranto31@hotmail.com / Universidad de Lima. Lima, Perú

Recepción: 31-5-2018 / Aceptación: 17-8-201

RESUMEN. Este artículo busca comparar dos metodologías de reconocimiento de expresiones faciales: Viola-Jones y Regression Based Facial Landmark Detection, adaptados para la detección de somnolencia, dando a conocer cuál de ellas es la más óptima y se adecúa mejor a las condiciones variables, considerando las restricciones de oclusión, rotación de rostro, iluminación. Tras un análisis cuantitativo bajo la matriz de confusión y poniendo a prueba las metodologías propuestas en diferentes situaciones, se realizó una comparación con los resultados obtenidos. Ocurren diferentes efectos por la falta de sueño como la disminución del tiempo de reacción, cansancio ocular, la visión se torna borrosa, menor concentración, etcétera; influyendo de manera directa en el incremento de accidentes de tráfico.

PALABRAS CLAVE: detección de somnolencia, reconocimiento facial, accidentes de tráfico, tasa de efectividad, tasa de reconocimiento, Viola-Jones, Regression Based Facial Landmark Detection, EAR

Adaptation and comparison of two facial recognition methods for driver drowsiness detection

ABSTRACT. This article seeks to make a comparison between two facial recognition methodologies: Viola-Jones and Regression Based Facial Landmark Detection and find the one that performs better in the face of different variable problems such as changing illumination, occlusion, face rotation, among others. The methodology that has a better rate of effectiveness for the detection of facial expressions is sought. Two facial detection methodologies will be adapted for the detection of drowsiness in drivers. After an analysis of advantages and disadvantages, a comparison of the results obtained will be shown. Different effects occur due to lack of sleep, such as decreased reaction time, eye fatigue, blurred vision, reduced concentration, etc.; These factors directly influence the increase in traffic accidents.

KEYWORDS: drowsiness detection, facial recognition, traffic accidents, effectiveness rate, recognition rate, Viola-Jones, Regression Based Facial Landmark Detection, EAR

1. INTRODUCCIÓN

La somnolencia en los conductores es uno de los factores que influyen en el índice de accidentes automovilísticos, ya que un conductor cansado o que se duerme durante su actividad no puede realizar maniobras evasivas para evitar las colisiones o despistes. Según la Dirección General de Circulación Terrestre del Ministerio de Transportes y Comunicaciones (MTC) del Perú, el año pasado se registraron 350 782 sujetos heridos y 32 107 muertos por imprudencia del conductor. Por estas cifras alarmantes de accidentes, no solo en el Perú sino en el mundo, sería necesaria la implementación de un sistema de detección de somnolencia efectivo en los vehículos para minimizar las probabilidades de que los accidentes de este tipo ocurran. En el caso de Lima, la congestión vehicular se agudiza por el número de vehículos que sigue aumentando año tras año, la infraestructura vial está llegando a su máxima capacidad, sobre todo en horas donde hay mayor número de conductores al volante —al final del día, cuando las personas están cansadas—, provocando en ellos estrés, aburrimiento, cansancio y sueño.

Diferentes efectos ocurren por la falta de sueño, como la disminución en el tiempo de reacción, cansancio ocular, la visión se torna borrosa, menor concentración, etcétera. Dichos factores influyen de manera directa en el incremento de accidentes de tráfico. Por ello, la ciudad de Lima es un lugar donde todos los autos deberían incorporar obligatoriamente un sistema de detección de somnolencia para disminuir el riesgo de que suceda algún accidente, por todas las razones propuestas anteriormente.

Existen varias técnicas para el reconocimiento de expresiones faciales, sin embargo, existen problemas en el momento en que se extraen las características faciales, tales como la oclusión, luminosidad, calidad de la imagen, la cercanía o lejanía del rostro, etcétera. Estos problemas entorpecen los resultados del reconocimiento y, por lo tanto, disminuye la tasa de efectividad. El problema radica en que se debe elegir una metodología que pueda superar todos estos problemas, o por lo menos los más importantes, y conseguir un porcentaje de reconocimiento bastante alto para llegar a ser preciso.

La presente investigación se enfocará en estudiar dos metodologías de reconocimiento facial: Viola-Jones y Regression Based Facial Landmark Detection (RBFLD) adaptadas al reconocimiento de somnolencia, y demostrar con cuál de ellas se obtiene una mayor tasa de efectividad.

2. METODOLOGÍA

2.1 Metodología Viola-Jones

Es un marco de trabajo desarrollado por los ingenieros Paul Viola y Michael Jones que utiliza una serie de algoritmos e ideas para una robusta y extremadamente rápida detección visual. El marco consta de tres principales contribuciones: un entrenador de clasificadores basado en el

algoritmo de aprendizaje AdaBoost, un algoritmo para la detección de objetos que utiliza la clasificación en “cascada”, y una nueva representación de la imagen denominada “imagen integral”. Inicialmente, para la detección del objeto se hace uso de la “imagen integral”, que permite que la evaluación de las características sea mucho más rápida; esta imagen se obtiene a través la realización de unas pocas operaciones por pixel que al finalizar permiten que la búsqueda de características en subregiones (de 24 por 24 píxeles), se transforme en una tarea de tiempo constante, sin importar su escala en la subregión o posición de la misma. Una vez obtenida la imagen el algoritmo la divide en subregiones de distintos tamaños y utiliza una serie de clasificadores, cada uno con un conjunto de características visuales, para discriminar si en la imagen se encuentra el objeto o no (Viola y Jones, 2004).

2.1.1 Implementación

La implementación del sistema de detección de somnolencia seguirá el diagrama de la figura 1. El proceso comienza con la captura de imagen *frame* por *frame*, hasta que la entrada deje de transmitirse; se detectará el rostro, a partir de ello se detectarán ojos y boca, y finalmente se enviará una alarma por parte del sistema si es que se detecta somnolencia en los participantes. Para la adaptación de Viola-Jones a la detección de somnolencia en imágenes se usó de base el código fuente proporcionado en el artículo “Facial Features Monitoring for Real Time Drowsiness Detection” (Manu, 2016).

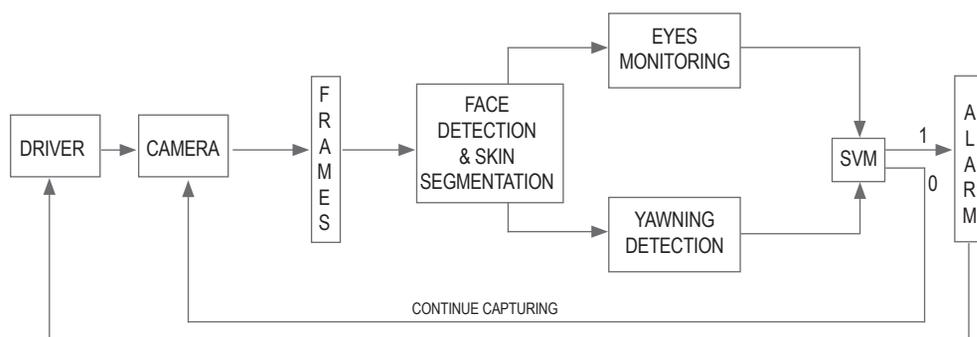


Figura 1. Diagrama de bloques de la implementación del algoritmo Viola Jones

Fuente: Manu, 2016

2.2 Metodología Regression Based Facial Landmark Detection

A cada *frame* del video se le aplicará el algoritmo de HOG para la detección de rostro. Esta técnica cuenta con ocurrencias de orientación de gradiente en porciones localizadas de una imagen. La imagen se divide en pequeñas regiones conectadas llamadas células y para los

pixeles dentro de cada célula se compila un histograma de direcciones de gradiente. La ventaja de este método es que la apariencia y la forma del objeto local dentro una imagen, en este caso *frame*, se puede describir mediante la distribución de gradientes de intensidad o direcciones de borde (Mallick, 2016).

Luego del reconocimiento del rostro, el paso siguiente es detectar los ojos. Para hacer esto posible se utilizan los Facial Landmark Detection, que son un subconjunto de puntos característicos en el rostro. Los puntos faciales más comunes son la punta de la nariz, los bordes de los ojos, el arco de la ceja, los lóbulos de las orejas, la boca, etcétera. En nuestro caso nos limitaremos a nuestras zonas de interés que son las zonas de los ojos, derecho e izquierdo.

Para detectar la somnolencia las partes fundamentales de interés son los ojos, tanto del lado izquierdo como del derecho. Estos están representados por coordenadas en un plano bidimensional tomando como inicio el ojo izquierdo y luego trabajando en sentido horario. Según Soukupová y Cech (2016), existe una relación entre el ancho y la altura de estas coordenadas; a partir de ello se puede derivar una ecuación que refleja esta relación conocida como Eye Aspect Ratio (EAR).

$$EAR = \frac{\|\rho_2 - \rho_6\| + \|\rho_3 - \rho_5\|}{2\|\rho_1 - \rho_4\|}$$

Donde: $\rho_1, \rho_2, \rho_3, \rho_4, \rho_5$ y ρ_6 son coordenadas del ojo

La implementación del sistema de detección de somnolencia seguirá el diagrama de la figura 2. El proceso comienza con la captura de imagen *frame* por *frame*, hasta que la entrada deje de transmitir; se detectará el rostro, luego se detectarán los ojos y finalmente se enviará una alarma por parte del sistema si es que se detecta somnolencia en los participantes. Para la adaptación de esta metodología para que detecte la somnolencia en imágenes se usó de base el código fuente proporcionado en el artículo “Real-Time Eye Blink Detection Using Facial Landmarks” (Soukupová y Cech, 2016).

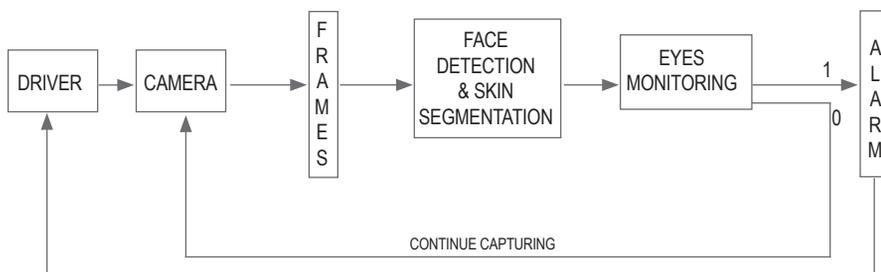


Figura 2. Diagrama de bloques de la implementación del algoritmo Regression Based Facial Landmark Detector
Elaboración propia

3. METODOLOGÍA DE LA EXPERIMENTACIÓN Y RESULTADOS

Para poder evaluar los resultados obtenidos con las metodologías de reconocimiento de somnolencia mencionadas anteriormente se ha analizado cada fotograma de cada uno de los videos proporcionados por la base de datos NTHU-Drowsy Driver Detection (NTHU-DDD) y también cada fotograma de los videos de nuestra propia base de datos.

Las dos metodologías muestran una alerta al momento de detectar la presencia de somnolencia y mediante una verificación manual contamos la cantidad de fotogramas que fueron detectados como somnolientos y la cantidad de los mismos que fueron clasificados como no somnolientos. Finalmente se usó un total de 32 videos proporcionados por la base de datos NTHU-DDD y otros 32 videos que grabamos para las pruebas.

Para la validación de nuestra propuesta de solución haremos uso de indicadores como: la precisión media (AAC) y la tasa de detección (DR) de acuerdo con las ecuaciones (1) y (2) respectivamente. Estas dos medidas nos permitirán evaluar la exactitud e indicar qué algoritmo propuesto es más aceptable en la detección de somnolencia en la cara de los conductores al momento de manejar su vehículo. Para ello haremos uso de una matriz de confusión donde:

- NTP: Sí tiene fatiga y sí ha detectado fatiga correctamente.
- NFN: No tuvo fatiga y sí ha detectado fatiga por error.
- NTN: No tuvo fatiga y no detectó fatiga.
- NFP: Sí tiene fatiga y no ha detectado fatiga por error.

$$AAC = \frac{(NTP + NTN)}{(NTP + NTN + NFP + NFN)} \dots\dots\dots(1)$$

$$DR = \frac{\text{Verdaderos Positivo}}{\text{Drowsy Frames}} = \frac{TP}{(NTP + NFN)} \dots\dots\dots(2)$$

Tabla 1

Matriz de confusión, resultados; situación 3: con oclusión y luminosidad alta. Método: Viola-Jones

Intento	Número de frames	Número de frames con fatiga			Número de frames sin fatiga	AAC (%)	DR (%)
		NTP	NFN	NTN	NFP		
1	1000	242	359	304	95	54,60	40
2	1000	194	423	345	38	53,90	31
3	1000	242	476	271	11	51,30	34
4	1000	321	243	423	13	74,40	57
5	1000	87	0	762	151	84,90	100
6	1000	285	167	515	33	80,00	63
7	1000	365	311	324	0	68,90	54
8	1000	115	68	664	153	77,90	63
Total	8000	1851	2047	3608	494	68,24	47

Elaboración propia

Se ha hecho el experimento sobre los videos de la base de datos NTHU-DDD, al igual que sobre videos que hemos grabado para nuestra propia base de datos en los que se ha pedido a los participantes que simulen síntomas de somnolencia, es decir, pestañear, bostezar, en pequeños intervalos de tiempo.

Para las pruebas realizadas se cuenta con cuatro tipos de escenario, cada uno de ellos con ocho videos de diferentes personas con características distintas en cuanto a etnia, color de piel, tamaño de ojo, etcétera. Los cuatro escenarios de prueba son los siguientes:

- Sin oclusión y luminosidad alta.
- Sin oclusión y luminosidad baja.
- Con oclusión y luminosidad alta.
- Con oclusión y luminosidad baja.

3.1 Resultados obtenidos con la base de datos NTHU-DDD

Contrastando los resultados obtenidos con las dos metodologías, con base en los videos de la base de datos NTHU-DDD, se muestran los siguientes resultados de los cuatro escenarios mencionados.

Tabla 2

Resultados de AAC finales de cada situación probada con la base de datos NTHU-DDD

	Metodología	
	Regression Based Facial Landmark Detection (%)	Viola-Jones (%)
Situación 1	88,65	59,38
Situación 2	74,88	81,70
Situación 3	85,25	73,30
Situación 4	73,88	75,90

Elaboración propia

En el tema de la precisión media (AAC) se puede apreciar que la metodología Regression Based Facial Landmark Detection (RBFLD) supera mínimamente a la metodología Viola-Jones en algunas situaciones, debido a la rotación de rostros que presentan los participantes en los videos.

Una de las ventajas que fue posible observar fue que el algoritmo de Viola-Jones se adecuaba mejor en las situaciones con luminosidad baja que el algoritmo RBFLD; por ello, en las situaciones 2 y 4 tuvo un mayor porcentaje que la otra metodología.

Una ventaja que se pudo observar del algoritmo RBFLD es que no tenía problemas en cuanto a la rotación del rostro, ya que en la mayoría de *frames* se detectó la zona de interés. Sin embargo, en la metodología de Viola-Jones se vieron dificultades con respecto a este ítem (rotación de rostro).

Tabla 3

Resultados de DR finales de cada situación probada con la base de datos NTHU-DDD

	Metodología	
	Regression Based Facial Landmark Detection (%)	Viola-Jones (%)
Situación 1	100,00	22,30
Situación 2	67,71	49,57
Situación 3	89,63	33,08
Situación 4	52,04	43,56

Elaboración propia

Al igual que la precisión media, la tasa de detección de RBFLD supera a Viola-Jones en este caso por un porcentaje alto en comparación a la del AAC. Esto sucedió por el mismo motivo, la rotación del rostro; Viola-Jones detectaba un número muy alto de “falsos negativos (NFN)” en todos los videos, mientras que RBFLD no. Se puede ver que ninguna de las dos metodologías tuvo un porcentaje de detección alto en la última situación probada, las dos se encuentran aproximadamente por el 50 % de detección lo cual en una situación real en que el conductor se quedara dormido, hay una probabilidad del 50 % de que no sea detectado y por ende no sea advertido de que se está quedando dormido, lo que podría ocasionar un accidente.

3.2 Resultados obtenidos con nuestra propia base de datos

Contrastando los resultados obtenidos de las dos metodologías con base en los videos que grabamos nosotros mismos, se muestran los siguientes resultados de los cuatro escenarios mencionados anteriormente.

Tabla 4
Resultados de AAC finales de cada situación con base de datos propia

	Metodología	
	Regression Based Facial Landmark Detection (%)	Viola-Jones (%)
Situación 1	87,03	60,35
Situación 2	86,23	67,10
Situación 3	83,56	68,24
Situación 4	88,92	66,14

Elaboración propia

En el tema de la precisión media (AAC) se puede apreciar que la metodología RBFLD supera en un 20 % a la metodología Viola-Jones en todas las situaciones. La metodología Viola-Jones, al momento de procesar los videos, daba como resultado varios frames como falsos negativos; por tal motivo se da la superioridad de RBFLD, ya que este detectó, la mayoría de las veces, la somnolencia cuando sí existía.

Cabe resaltar que frente al experimento de la base de datos NTHU-DDD, Viola Jones no tuvo problemas al detectar los ojos del participante ya que en esta ocasión no se encontró una rotación del rostro del mismo y la cámara se encontraba frontalmente al rostro del participante.

Una ventaja que se pudo observar del algoritmo RBFLD es que no tenía problemas en cuanto a la rotación del rostro, ya que en la mayoría de *frames* se detectó la zona de interés que en este caso son los ojos.

Tabla 5
Resultados de DR finales de cada situación con base de datos propia

	Metodología	
	Regression Based Facial Landmark Detection (%)	Viola-Jones (%)
Situación 1	95,00	53,45
Situación 2	95,00	47,00
Situación 3	79,00	47,00
Situación 4	95,00	42,00

Elaboración propia

Al igual que la precisión media (AAC), la tasa de detección de RBFLD supera a Viola-Jones, en este caso por un porcentaje alto en comparación al del AAC, al igual que en el experimento con la base de datos NTHU-DDD. Esto sucedió porque Viola-Jones detectaba un número muy alto de falsos negativos en todos los videos, al igual que en el otro experimento; en cambio, el RBFLD no. Se puede ver esta vez que el RBFLD tuvo un porcentaje de detección alto en todas las situaciones excepto en la tercera. Esto se dio porque se detectaba la zona del párpado inferior/superior en vez de los ojos debido al uso de lentes, provocando que, si la persona estaba fatigada, no detectara el algoritmo ya que tomaba la apertura del párpado.

4. CONCLUSIONES

Para realizar nuestra comparación de metodologías usamos una base de datos NTHU-DDD brindada gracias a la universidad National Tsing Hua, de Hsinchu, Taiwán, donde encontramos diferentes videos de personas de etnias distintas simulando que estaban somnolientas en cuatro diferentes situaciones: con luminosidad alta y sin oclusión, con luminosidad baja y sin oclusión, con luminosidad alta y con oclusión, y con luminosidad baja y con oclusión. Con respecto a la precisión media (ACC), con el algoritmo de RBFLD se obtuvo un ACC, en las diferentes situaciones, de 88,65 %, 74,88 %, 85,25 % y 73,88 % respectivamente, en comparación con el algoritmo de Viola-Jones que obtuvo porcentajes de ACC de 59,38 %, 81,70 %, 73,30 % y 75,90 % respectivamente. Estos resultados se dieron ya que el algoritmo

de Viola-Jones, en algunas situaciones, no detectaba correctamente los *frames* debido a la rotación del rostro que el participante presentaba. Sin embargo, este mismo algoritmo respondía mejor a las situaciones de luminosidad baja. Con respecto a la tasa de detección de somnolencia (DR), en el algoritmo RBFLD se obtuvo un DR, en las diferentes situaciones, de 100 %, 67,71 %, 89,63 % y 52,04 % respectivamente, en comparación con el algoritmo de Viola-Jones, que obtuvo porcentajes de 22,30 %, 49,57 %, 33,08 % y 43,56 % respectivamente. Estos resultados se dieron por el mismo motivo mencionado anteriormente, la rotación del rostro fue un problema para el algoritmo de Viola-Jones, captando un número alto de falsos negativos.

Debido a que se dieron estos resultados, decidimos profundizar nuestra validación probando ambos algoritmos en una base de datos con rostros latinos realizada por nosotros mismos con ayuda de nuestros compañeros de la universidad, para luego comparar resultados. Se plantearon las mismas situaciones definidas inicialmente: con luminosidad alta y sin oclusión, con luminosidad baja y sin oclusión, con luminosidad alta y con oclusión, y con luminosidad baja y con oclusión. Con respecto a la precisión media (ACC), con el algoritmo RBFLD se obtuvo un ACC, en las diferentes situaciones, de 87,03 %, 86,23 %, 83,56 % y 88,92 % respectivamente, en comparación con el algoritmo de Viola-Jones que obtuvo porcentajes de ACC de 60,35 %, 67,10 %, 68,24 % y 66,14 % respectivamente. Estos resultados se dieron porque el algoritmo de Viola-Jones, al procesar los videos, captaba varios *frames* como falsos negativos a pesar de que el algoritmo no tuviera el problema de rotación de rostro y se detectaban los ojos de los participantes. Con respecto a la tasa de detección de somnolencia (DR), en el algoritmo RBFLD se obtuvo un DR, en las diferentes situaciones, de 95 %, 95 %, 79 % y 95 % respectivamente, en comparación con el algoritmo de Viola-Jones que obtuvo porcentajes de 53,45 %, 47 %, 47 % y 42 % respectivamente. Estos resultados se dieron por la misma razón mencionada anteriormente: Viola-Jones, al procesar los videos, captaba varios *frames* como falsos negativos a pesar de que el algoritmo no tenía el problema de rotación de rostro y se detectaban los ojos de los participantes.

Pensamos que en el caso de la metodología Viola-Jones, la inclusión de la detección del bostezo para detectar la somnolencia, en vez de acertar mejor, empeoraba la precisión de la detección de la somnolencia. En cambio, la metodología RBFLD, al concentrarse solo en la apertura de los ojos, obtenía un mejor resultado, como se muestra en los porcentajes. Finalmente, de acuerdo con los resultados obtenidos, en la mayoría de los escenarios la metodología RBFLD detecta mejor la somnolencia que Viola-Jones, superando las condiciones variables de oclusión, luminosidad y rotación del rostro.

Con base en los trabajos realizados en este campo de investigación, se puede apreciar el aporte en el tema de validación con una cantidad mayor de *frames* que, a su vez, refleja una mayor cantidad de videos. Además, elaboramos nuestra propia base de datos con el objetivo de analizar cómo actúan los algoritmos frente a rostros latinos, específicamente peruanos.

Por último, en la adaptación de Viola-Jones para detectar la somnolencia se usó de base el código fuente proporcionado en el artículo “Facial Features Monitoring for Real Time Drowsiness Detection” (Manu, 2016). Se tuvo que modificar en algunos aspectos, por ejemplo, para que aceptara como *input* videos en vez de utilizar una cámara en tiempo real; se agregó un contador de *frames* para automatizar la validación y ver en qué *frame* del video se perciben los distintos escenarios de la matriz de confusión (verdaderos positivos, verdaderos negativos, falsos positivos, falsos negativos). En el caso del algoritmo RBFLD, se usó de base el código fuente proporcionado por Adrian Rosebrock y se modificaron ciertos parámetros, por ejemplo, para que aceptara videos como input en vez de utilizar la cámara en tiempo real, se agregó un contador de *frames* para automatizar la validación y se agregó al código la fórmula de la apertura de los ojos proporcionada en el artículo “Real-Time Eye Blink Detection Using Facial Landmarks” (Soukupová y Cech, 2016).

REFERENCIAS

- Manu, B. N. (noviembre de 2016). Facial features monitoring for real time drowsiness detection. En: *Innovations in Information Technology (IIT), 2016 12th International Conference*, pp. 1-4. IEEE.
- Rosenbrock, A.
- Soukupová, T., y Cech, J. (Febrero de 2016). Real-Time Eye Blink Detection Using Facial Landmarks. En: *21st Computer Vision Winter Workshop*.
- Viola, P., y Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision* 57(2), pp. 137–154.

BIBLIOGRAFÍA

- Aidman, E., Chadunow, C., Johnson, K., y Reece, J. (2015). Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. *Accident Analysis & Prevention* 81, pp. 8-13.
- Banco Mundial. (2017). *Población urbana (% del total)*. Recuperado de <http://datos.bancomundial.org/indicador/SP.URB.TOTL.IN.ZS?end=2015&start=1985&view=chart>
- Bazrafkan, S., Nedelcu, T., Filipczuk, P., y Corcoran, P. (2017). Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods. *IEEE International Conference on Consumer Electronics (ICCE)*.

- Cheng, B., Zhang, W., y Lin, Y. (2012). Driver drowsiness recognition based on computer vision technology. *Tsinghua Science and Technology* 17, pp. 354-362.
- Choy, M., y Chang, G. (2014). *Medidas macroprudenciales aplicadas en el Perú*. Lima: Banco Central de Reserva del Perú. Recuperado de <http://www.bcrp.gob.pe/docs/Publicaciones/Documentos-de-Trabajo/2014/documento-de-trabajo-07-2014.pdf>
- Dwivedi, K., Biswaranjan, K., y Sethi, A. (2014). Drowsy driver detection using representation learning. *Advance Computing Conference (IACC)*, IEEE International.
- Flores Calero, M. J. (2009). *Sistema avanzado de asistencia a la conducción mediante visión por computador para la detección de la somnolencia*. Tesis doctoral. Universidad Carlos III, Madrid, España.
- Hu, S., y Zheng, G. (2009). Driver drowsiness detection with eyelid related parameters by Support Vector Machine. *Expert Systems with Applications*, pp. 7651-7658.
- Hwang, K.-A., y Yang, C.-H. (2009). Attentiveness assessment in learning based on fuzzy logic analysis. *Expert Systems with Applications* 36, pp. 6261-6265.
- Julián García, R. A., y Crespín Luis, J. C. (2014). *Sistema detector de somnolencia en secuencias de video de conductores manejando usando visión computacional*. Tesis de pregrado. Universidad Nacional de Trujillo.
- Lee, B. G., Jung, S. J., y Chung, W.-Y. (2011). Real-time physiological and vision monitoring of vehicle driver for non-intrusive drowsiness detection. *IET Communications* 5, pp. 2461-2469.
- Pereyra, P. A. (S. f.). *Reconocimiento facial*. Tesis de pregrado. Universidad de Buenos Aires.
- Riveros, O. A. (2008). *Análisis y detección de características faciales usando aprendizaje estadístico*. Tesis de posgrado. Universidad de Chile.
- Rosebrock, A. (2018). Face recognition with OpenCV, Python, and deep learning. Recuperado de <https://www.pyimagesearch.com/2018/06/18/-face-recognition-with-opencv-python-and-deep-learning/>
- Triantafillou, D., y Tefas, A. (2016). Face detection based on deep convolutional neural networks exploiting incremental facial part learning. *23rd International Conference on Pattern Recognition (ICPR)*.
- Wang, X., y Xu, C. (2016). Driver drowsiness detection based on non-intrusive metrics considering individual specifics. *Accident Analysis & Prevention* 95, pp. 350-357.
- Zhang, T., Zheng, W., Cui, Z., Zong, Y., Yan, J., y Yan, K. (2016). A Deep Neural Network-Driven Feature Learning Method for Multi-view Facial Expression Recognition. *IEEE Transactions on Multimedia* 18, pp. 2528-2539.