

Identificación de Conductores Somnolientos usando aprendizaje automático

Esther Durán Reyes¹, Gustavo Rafael Guzmán Loreto², André Ramos Moreno², Paolo Giovanni Ceretti Gomez², Christian Agustín Tavera Rocha² y Rafael Guzmán Cabrera¹

¹División de Ingenierías Campus Irapuato-Salamanca Universidad de Guanajuato

²División de Ciencias e Ingenierías Campus León Universidad de Guanajuato
e.duranreyes@ugto.mx¹, gr.guzmanloreto@ugto.mx², a.ramosmoreno@ugto.mx², pg.cerettigomez@ugto.mx², ca.taverarocha@ugto.mx², guzmanc@ugto.mx¹

Resumen

Todos los días ocurren accidentes automovilísticos. Muchos de estos accidentes son provocados por que los conductores se quedan dormidos o están cansados. Si se pudiera contar con un sistema que permitirá activar una alerta al momento de detectar que el conductor se está quedando dormido, el número de accidentes se reduciría. En el presente trabajo se realiza la clasificación automática de imágenes de conductores somnolientos. Utilizamos una base de datos manualmente etiquetada, técnicas de aprendizaje automático y distintos métodos de aprendizaje. se desarrolló una interfaz gráfica para preprocesar las imágenes. El mejor resultado obtenido es del orden del 95% de precisión usando regresión logística, lo cual habla de la viabilidad del método implementado. se utilizaron distintas métricas de evaluación utilizadas en el estado del arte.

Palabras Clave: Aprendizaje automático, Identificación de conductor somnoliento, procesamiento de imágenes.

Introducción.

Según datos de la Organización Mundial de la Salud (WHO, por sus siglas en inglés) en el Global status report on road safety 2015 [1], los accidentes automovilísticos ascienden a 1,250,000 percances cada año, de los cuales se estima que entre el 15-20% están relacionados directamente con conductores somnolientos [2]. No todos los accidentes de tránsito que ocurren lo hacen por efectos del alcohol en los conductores, ya que puede haber factores externos, como fallas mecánicas, exceso de velocidad, el tráfico, condiciones climatológicas adversas como la lluvia, nieve y viento que complican que el conductor promedio disminuya su desempeño al manejar. En el presente trabajo llevamos a cabo la clasificación automática de imágenes de conductores somnolientos. Para hacer este trabajo nos basamos en un conjunto de imágenes que fueron manualmente etiquetados y se describen más adelante. Con la finalidad de ir introduciéndonos en tema, comenzaremos definiendo lo que es la Somnolencia: Es la sensación de pesades de los sentidos que surgen a partir del sueño, esto produce en una persona la disminución de la capacidad de vigilia, percepción y control. Es considerada un trastorno del sueño ya que se origina por distintas enfermedades como lo son la diabetes, narcolepsia y depresión, entre otras.

La mayor causa de la somnolencia es la alteración del ritmo circadiano, este es el cambio de las funciones fisiológicas del ser humano tal como el metabolismo, respiración, contracción muscular, ciclos del sueño, temperatura del cuerpo, etc. También tenemos que hablar del micro sueño, este puede durar entre 2 y 4 segundos, la principal causa de los micros sueños es la privación del sueño y la falta de un descanso adecuado. Por otro lado, un sistema es un conjunto de partes o elementos organizados y relacionados que interactúan entre sí para lograr un objetivo. Está compuesto por entradas donde se recibe datos o información y las salidas por donde se provee la información. Y en nuestro caso particular, un sistema de procesamiento de imágenes: Este sistema se define como los procesos de obtención e interpretación de las imágenes, los procesos constan de la adquisición, preprocesado, segmentación, representación y descripción y reconocimiento e interpretación.

Este tema de investigación ya ha sido estudiado por investigadores en el mundo, por ejemplo en [3] se utilizan características Gabor aplicadas sobre la cara para detectar la fatiga. Primero se localiza la cara y se obtienen las características Gabor a diferentes escalas y orientaciones. Las características de la misma escala se fusionan para reducir la dimensionalidad y finalmente, se utiliza el algoritmo de AdaBoost para extraer las características más críticas y construir un clasificador robusto que detecte la somnolencia. Este método es robusto ante un amplio rango de sujetos expuestos a diferentes posturas y condiciones de iluminación.

En [4] se usan los modelos de mezclas de Gaussianas para modelar estadísticas del comportamiento normal de la duración del cierre de los ojos (ECD), y de la frecuencia del cierre de los ojos (FEC), para cada una de las personas e identificar comportamientos anómalos en el cierre de los ojos.

En [5] se utilizan Redes Neuronales Artificiales (ANN) para analizar parámetros del vehículo y datos de cierre de ojos para inferir la fatiga o somnolencia del conductor.

En [6] para mejorar la exactitud de la medida de somnolencia se combina la desviación estándar de la dirección de la mirada, la orientación de la cara, los diámetros de las pupilas y el pulso cardíaco.

El resto del trabajo se encuentra estructurado de la siguiente manera: en la sección Procedimiento se describe el proceso que se llevó a cabo, desde la selección de los elementos necesarios para realizar el trabajo, desde la base de datos utilizada hasta la descripción del modelo entrenado, pasando por los algoritmos y escenarios en que se realizó la clasificación de imágenes, además de una descripción general de cómo se llevó a cabo el proceso. Posteriormente, en la sección de Resultados, se expone cómo es que la conjunción de los elementos ya mencionados (base de datos, algoritmos, modelo, escenarios de clasificación) varía en las diferentes métricas de evaluación. En la sección de Conclusión se discuten los resultados y se define qué algoritmo fue óptimo bajo qué escenario de clasificación y en qué métricas exactamente, con las consideraciones necesarias.

Metodología.

En la figura 1 se muestra una ilustración esquemática que muestra un resumen de la metodología implementada en el presente trabajo. A continuación, presentamos una breve descripción de cada una de las etapas.

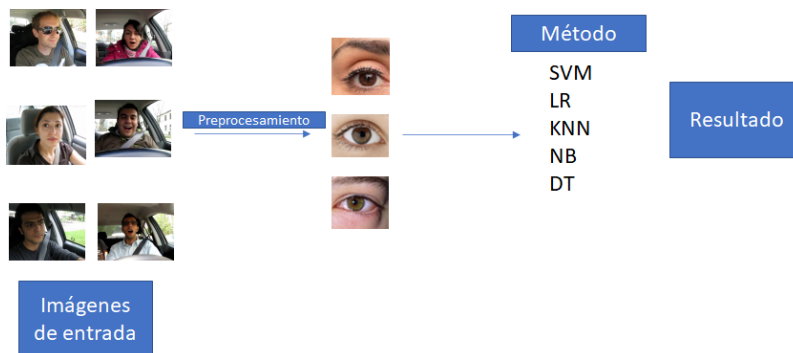


Figura 1. Metodología implementada para la clasificación de imágenes de conductores somnolientos.

Las imágenes utilizadas en el presente trabajo corresponden a la base de datos: 'driver drowsiness using keras' [7], misma que cuenta con cuatro clases diferentes de imágenes, precisas para saber de entrada si un ser humano se encuentra bajo la influencia del sueño; Open y Closed para los ojos del conductor, y decidir que, de coincidir más con los datos de la clase Closed es muy probable que nos encontramos con un conductor en estado de somnolencia y que es potencialmente peligroso. Sin embargo, la base de datos va más allá y nos brinda datos para hacer un análisis más completo del rostro del conductor analizado mediante las categorías de imágenes yawn y no_yawn que son para analizar la boca del conductor y saber si está

bostezando, lo cual, si bien no implicaría que ya está dormido, si diría que tiene sueño o está cansado. Cada una de las clases de imágenes contenidas en esta base de datos cuenta con un aproximado de 725 imágenes, lo cual nos permite realizar un análisis fiable y completo de la problemática, con un total de 2900 imágenes. El número de imágenes, así como la clase a la que pertenece cada una de ellas se muestra en la tabla 1.

Clase	Total de imágenes	Entrenamiento 80%	Prueba 20%
Open	726	581	145
Closed	726	581	145
yawn	723	578	145
no_yawn	725	580	145
Total	2900	2320	580

Tabla 1. Características de la base de datos. Open, Closed (para los ojos de la persona) y yawn, no_yawn (para la boca de las personas)

A continuación, en las figuras 2 a 5, se muestran algunas de las imágenes correspondientes a cada una de las clases:

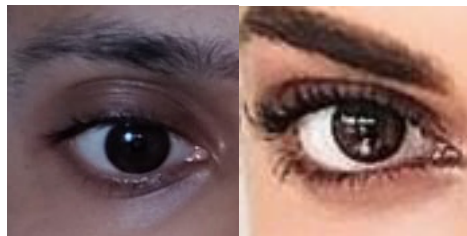


Figura 2. Imágenes de la clase Open

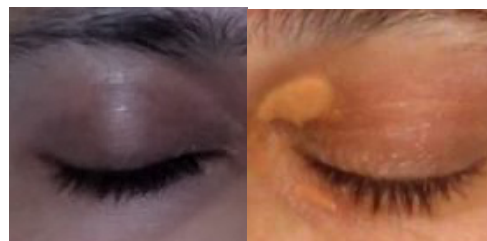


Figura 3. Imágenes de la clase Closed



Figura 4. Imágenes de la clase yawn

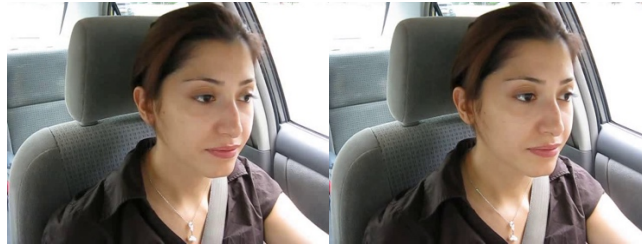


Figura 5. Imágenes de la clase no_yawn

A continuación, se presenta una breve descripción de los métodos de aprendizaje automático utilizados en el desarrollo del presente trabajo:

KNN: Es un clasificador de aprendizaje supervisado no paramétrico, que utiliza la proximidad (k vecinos más cercanos) para hacer clasificaciones o predicciones sobre la agrupación de un punto de datos individual [14].

Tree: Es un diagrama de flujo que empieza con una idea principal y luego se ramifica según las consecuencias de las decisiones. Se denomina “árbol de decisiones” por su semejanza con un árbol con muchas ramas [15].

SVM: Support vector machine (SVM) es un algoritmo de aprendizaje supervisado que se utiliza en muchos problemas de clasificación y regresión, incluidas aplicaciones médicas de procesamiento de señales, procesamiento del lenguaje natural y reconocimiento de imágenes y voz.

Naive Bayes: Es una técnica de clasificación basada en el Teorema de Bayes con un supuesto de independencia entre predictores; asume que la presencia de una característica particular en una clase no está relacionada con la presencia de ninguna otra característica.

Logistic Regression: La regresión logística es un método estadístico conocido que se usa para determinar la contribución de varios factores a un par de resultados. La implementación de Microsoft usa una red neuronal modificada para modelar las relaciones entre las entradas y los resultados.

Las métricas de evaluación sirven para cuantificar y evaluar aspectos de tendencia, comportamientos y resultados. Además, valoran el rendimiento de un modelo de aprendizaje automático, que es un componente integral en la ciencia de los datos. Para medir el desempeño de estos métodos de aprendizaje automático a la hora de realizar la clasificación automática de imágenes, se utilizaron las siguientes métricas de evaluación:

En la tabla 2 se muestra una matriz de confusión, a partir de esta, se pueden generar varias métricas para evaluar el desempeño del clasificador.

Tabla 2. Matriz de confusión.

	Clase positiva real	Clase negativa real
Clase positiva predicha	Verdadero positivo (<i>tp</i>)	Falso positivo (<i>fp</i>)
Clase negativa predicha	Falso negativo (<i>fn</i>)	Verdadero negativo (<i>tn</i>)

Precisión

La precisión indica con qué frecuencia es correcto y se obtiene a partir de la Ec. (2). Esta métrica representa el número de verdaderos positivos que son realmente positivos en comparación con el número total de valores positivos predichos.

$$Precision = \frac{tp}{tp+fp}$$

Recall (exhaustividad/sensibilidad). La métrica de exhaustividad muestra la cantidad de verdaderos positivos que el modelo ha clasificado en función del número total de valores positivos, como se muestra en la siguiente ecuación:

$$Recall = \frac{tp}{tp+fn}$$

Medida F. Esta métrica es la combinación de las métricas de precisión y recall, y sirve de compromiso entre ellas. La mejor puntuación es igual a 1 y la peor 0.

$$F1 = 2 * \frac{precision*recall}{precision+recall}$$

Accuracy. Es el porcentaje total de elementos clasificados correctamente. Es la medida más directa de la calidad de los clasificadores. Es un valor entre 0 y 1. Cuanto más alto, mejor, Accuracy es una buena medida cuando las clases de variables de destino en los datos están casi equilibradas.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

AUC. Se calcula como el área bajo una curva ROC (Receiver Operator Characteristic, característica de operador receptor) y es una representación escalar del rendimiento esperado de un clasificador. El AUC se sitúa siempre entre 0 y 1, y cuanto más alto es el valor, mejor es el clasificador. Una curva ROC diagonal entre las coordenadas (0,0) y (1,1) representa un clasificador aleatorio y tiene un AUC de 0,5. Así pues, un clasificador realista no tendrá un AUC de menos de 0,5.

$$1 - ESPECIFICIDAD = \frac{FP}{VN + FP}$$

Una vez descritos los métodos de aprendizaje automático utilizados, así como las métricas de evaluación, en la siguiente sección se presentan los resultados obtenidos.

Resultados

En la tabla 3 se muestran los resultados obtenidos. En la primera columna podemos identificar los distintos métodos de aprendizaje automático utilizados para llevar a cabo la clasificación automática de imágenes, mientras que, en la primer fina se observan las distantes métricas de evaluación descritas en el apartado anterior. En este caso las imágenes utilizadas corresponden al tipo Yawn-Noyawn.

Tabla 3. Resultados obtenidos

	YAWN-NOYAWN				
kNN	0.947179854	0.85980663	0.859262954	0.86530495	0.85980663
Tree	0.752898364	0.754834254	0.754833903	0.754840033	0.754834254
SVM	0.843893738	0.738259669	0.73690245	0.743130982	0.738259669
Naive Bayes	0.787190347	0.694060773	0.679360422	0.737124579	0.694060773
Logistic Regression	0.982833977	0.944751381	0.944719109	0.945730926	0.944751381

En la gráfica de la figura 6 se muestran las métricas de evaluación obtenidas para los distintos métodos de clasificación automática de imágenes. Podemos observar que el mejor desempeño se tiene con el método

regresión logística, mientras que el peor desempeño lo tiene Naive Bayes. El orden de las métricas, de izquierda a derecha es este: AUC, CA, F1, PRECISION y RECALL.

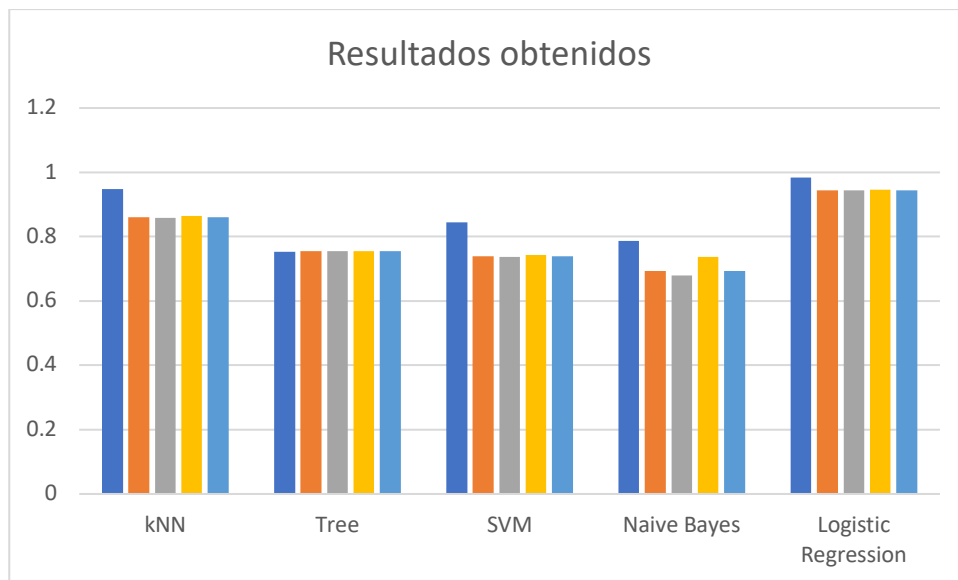


Figura 6. Tipos de Clasificadores

Conclusiones

Cada algoritmo parte de un principio matemático distinto, por lo que, dadas las características de las imágenes que se deseen clasificar habrá diferentes resultados en las comparaciones de algoritmos. Dicho sea, los modelos de aprendizaje implementados, pueden ser utilizados para otros procesos que no sean necesariamente de clasificación; regresión y predicción, por ejemplo. El punto aquí es que, al variar el qué haremos con nuestro modelo, tendremos también diferentes resultados en la comparativa de algoritmos.

Entonces cabe también concluir, el trabajo de clasificación de imágenes para la detección de conductores con rasgos de somnolencia se óptimamente realizado en un escenario de clasificación de conjuntos de entrenamiento y prueba y con SMV o LR cómo algoritmos, considerando también la red utilizada.

En la continuidad del trabajo y con objetivo de mejora, se podrían comparar los resultados aquí expuestos en similitud de condiciones con otro modelo. Además, se puede desarrollar una interfaz electrónica utiliza una placa Arduino y módulos de captura/envío de datos, y se monitorea con alertas visuales si se encuentran los ojos cerrados, fuera de la región de captura y a una distancia adecuada para el manejo del vehículo.

Referencias

- WHO, Global status report on road safety 2015. World Health Organization, 2015
- Home, J. & Reyner, Louise. (1999). Vehicle accidents related to sleep: A review. Occupational and environmental medicine. 56. 289-94. 10.1136/oem.56.5.289.
- Li, H. A., Fan, J., Zhang, J., Li, Z., He, D., Si, M., & Zhang, Y. (2021). Facial image segmentation based on Gabor filter. Mathematical Problems in Engineering, 2021.

- D'Orazio, D. J., & Loeb, A. (2018). Repeated imaging of massive black hole binary orbits with millimeter interferometry: Measuring black hole masses and the hubble constant. *The Astrophysical Journal*, 863(2), 185.
- Eskandarian, S., Bahrami, P., & Kazemi, P. (2017). A comprehensive data mining approach to estimate the rate of penetration: Application of neural network, rule based models and feature ranking. *Journal of Petroleum Science and Engineering*, 156, 605-615.
- Miyaji, M. (2014). Method of drowsy state detection for driver monitoring function. *International Journal of Information and Electronics Engineering*, 4(4), 264.
<https://www.kaggle.com/adinishad/driver-drowsiness-using-keras/data>
- Dua, M., Shakshi, Singla, R. et al. Deep CNN models-based ensemble approach to driver drowsiness detection. *Neural Comput & Applic* 33, 3155–3168 (2021). <https://doi.org/10.1007/s00521-020-05209-7>
- Wendy Morales Castro. (2019). Escenarios de clasificación. En *Detección automática de patrones en imágenes de tuberculosis pulmonar*. (81). Salamanca, Gto, Mex: Universidad de Guanajuato.
- Guyon, I., Weston, J., Barnhill, S. and Vapnik, V. (2001) Gene selection for cancer classification using support vector machines. *Machine Learning*, 46, 389–422
- Leif E. Peterson (2009) K-nearest neighbor. *Scholarpedia*, 4(2):1883.
- ALVAREZ VÁSQUEZ KATHERINE YESSÉNIA. (11 de 12 de 2016). PROPUESTA DE DESARROLLO DE UN SISTEMA PARA LA PREVENCIÓN DE. Obtenido de <http://repositorio.ug.edu.ec/bitstream/redug/56191/1/tesis%20para%20Universidad.pdf>
- CEA, F. (2016). Estudio somnolencia en la conducción. Obtenido de <https://www.fundacioncea.es/actualidad/estudios-fundacion/66-estudio-fundacion-cea-estudio-somnolencia-en-la-conduccion>
- Algoritmo de k vecinos más cercanos. (s.f.). Obtenido de: <https://www.ibm.com/mx/es/topics/knn#:~:text=El%20algoritmo%20de%20k%20vecinos%20m%C3%A1s%20cercanos%2C%20tambi%C3%A9n%20conocido%20como,un%20punto%20de%20datos%20individual.>
- Cómo hacer y usar árboles de decisiones. (s.f.). Obtenido de: <https://lucidspark.com/es/blog/como-hacer-arboles-de-decisiones>