

Detección de Motocicletas y del uso del casco de seguridad utilizando un algoritmo que emplea técnicas de procesamiento de imágenes y modelos de inteligencia artificial

Ing. Mario Andrés Varon
Universidad de los Andes
Candidato de Maestría
Bogota, Colombia
ma-varon@uniandes.edu.co

José Tiberio Hernández PhD
Universidad de los Andes
Profesor de Cátedra
Bogotá, Colombia
jhernand@uniandes.edu.co

Resumen-Abstract:

En Colombia los motociclistas son las primeras víctimas de la accidentalidad vial. Desde el 2001 hasta el 2014, se cuentan cerca de 28.000 fallecidos. Cerca de la mitad de dichos fallecimientos se producen como consecuencia de la falta de uso de elementos de protección pasiva (casco de seguridad). Para abordar este problema, se propone un algoritmo que utiliza técnicas de procesamiento de imágenes en conjunto con modelos de inteligencia artificial para la detección de motos y el uso de casco por parte de quienes las conducen. La precisión en la clasificación de motos alcanza el 97,14%, mientras que esa medida en la clasificación de cascos es del 85,29%.

Palabras Claves: Detección de Motos, Detección de Casco de Seguridad, Procesamiento de Imágenes, Inteligencia Artificial, Deep Learning, Boosted Trees, Umbralización por momentos, Umbralización por máxima entropía, Sobel, Llenado Morfológico, HSB stack.

1. Introducción

En Colombia los motociclistas son las primeras víctimas de la accidentalidad vial (Instituto Javeriano de Salud Pública, 2015). En el año 2001, representaban cerca del 24% de la mortalidad total producida en colisiones de tráfico. Dicho valor se incrementó hasta el 45% en el 2014. En tal lapso, se contabilizaron cerca de 28.000 motociclistas fallecidos, de los cuales el 86,6% eran hombres. La circunstancia explicativa del aumento de la mortalidad relativa y absoluta de los motociclistas se debe en primera circunstancia, al

crecimiento desbordado del parque automotor, con un incremento del 445% en 14 años (El Espectador, 2015) y en segunda medida a la falta de uso de elementos de protección pasiva, como el casco de seguridad.

Para el año 2013, se presentaron 6.579 defunciones como resultado de accidentes de tránsito a lo largo del territorio nacional. Esto es una tasa de 13,76 por cada 100.000 Habitantes. (Dirección de Epidemiología y Demografía, Ministerio de Salud, 2015). La alta tasa de mortalidad dentro de este grupo poblacional subyace en la falta de uso elementos de seguridad pasiva. Mientras en un vehículo la energía de la

colisión de tráfico la recibe en mayor parte el chasis de este, en un motociclista, es su cuerpo quién se encarga de recibir y disipar la energía que produce el choque. De ahí deriva la importancia de utilizar los elementos de protección en la conducción de motocicletas, para mitigar sustancialmente el impacto de dichas colisiones. En especial, el uso del casco de seguridad reglamentario, disminuye el riesgo de mortalidad en este grupo poblacional hasta en un 39%, mientras que reduce también la posibilidad de sufrir trauma craneo encefálico hasta en un 72%. (Loaiza, 2012).

No existe evidencia directa que permita inferir la proporción de la población de motociclistas que hace uso de sistemas de protección pasiva. Por el contrario, dichos valores se han obtenido del estudio indirecto del análisis forense del occiso con posterioridad a la ocurrencia del accidente de tránsito. Con dicha metodología se concluyó que existe ausencia del casco u omisión de la normatividad técnica en el que se porta, en cerca de seis (6) de cada diez motociclistas fallecidos (10) (Valbuena, 2012). Así, resulta fundamental desarrollar una herramienta que permita detectar de forma automática y no intrusiva, el uso de del casco de seguridad por parte de los motociclistas con el fin de desarrollar e implementar políticas públicas, por parte de las autoridades competentes, que permitan mitigar los altos costos que produce la accidentalidad vial al interior del país.

Se presenta en este artículo un algoritmo que incluye el uso de técnicas de procesamiento de imágenes conjunto con el despliegue de modelos de inteligencia artificial, que permiten agendar este problema de forma precisa, ágil y a costo moderado con el fin de solventar un problema que es común en países en vía de desarrollo.

2. Definición del Problema

Para atender las problemáticas de seguridad vial, que se mencionan con anterioridad, se requiere la detección de forma precisa, ágil y a bajo costo de las motocicletas dentro del flujo vehicular, así como la detección del casco de seguridad por parte de aquellos que la conducen.

Para tal fin, se han creado los Sistemas Inteligentes de Transporte (ITS). Estos son aplicaciones avanzadas que combinan sistemas electrónicos, de comunicaciones, de computadores y sensores. Dichos ITS integran vehículos, personas e información de la vía con estrategias de gestión del tráfico para proveer información en tiempo real para

aumentar la seguridad, la eficiencia y el confort en los sistemas de tráfico tanto públicos, como privados. (Roess, Prassas, & McShane, 2004).

En especial, en el contexto de desarrollo de los distintos ITS, se ha permitido realizar la detección de motocicletas por una gran variedad de tecnologías. Entre ellas, los radares, los sensores LIDAR, los sensores de sonido y los sensores ópticos. (Amir, Likun, & Tang, 2013). Dichas tecnologías para la detección de motocicletas se pueden dividir en dos grupos (Herbert, 2000). Aquellas basadas en sensores activos (i.e los sensores aquí descritos identifican los objetos con la reflexión de la señal emitida por ellos.) y de otro lado están los sensores pasivos (i.e los sensores adquieren los datos sin emitir ninguna señal).

Se debe mencionar que la detección de motocicletas y del casco de seguridad, es una tarea bastante ardua en la medida en que ambos objetos presentan una alta variabilidad en su forma, color y tamaño. Se añade a esto, ambientes de tránsito altamente congestionados, cambios en la iluminación y la interacción aleatoria entre los diferentes actores viales que participan en el día a día del tráfico de una ciudad. Estos pueden generar occlusión sobre cualquiera de los objetos descritos. De otro lado, las distintas tecnologías para la detección y reconocimiento de motocicletas, se enfrentan a dos problemas de gran magnitud: La capacidad de cómputo -ya que la mayoría de ellos, se encuentran desplegados en la vía- y el grado de confiabilidad para la detección de ellas.

En la medida en que el algoritmo propuesto, involucra el despliegue de sensores pasivos, por su alta precisión y su bajo costo, se procede a construir el estado del arte vigente, para la detección y clasificación de motocicletas y el reconocimiento del uso del casco de seguridad en quien la conduce.

3. Estado del Arte

a. Detección, Reconocimiento y Clasificación de Motocicletas.

Tal y como se describe en (Amir, Likun, & Tang, 2013) existen una gran variedad de técnicas para la detección, el reconocimiento y clasificación de Motocicletas en el tráfico vehicular.

Todas aquellas basadas en técnicas de visión que se han propuesto para la detección de motocicletas pueden encontrarse en (Chung-Cheng, Min-Yu, & Hung-Tsung, 2007) (Kanhare, Birchfield, Sarasua,

& Khoemi, 2010) (Meeselodi, Modena, & Cattoni, 2007) (Min-Yu, Chung-Cheng, Hung-Tsung, & Hong, 2008) (Nguyen & Le, 2008) (Wen-Chung & Kuo-Jung, 2010) (Zezhi, Ellis, & Velastin, 2011). En la mayoría de los casos que se describen aquí, la cámara está estática, con lo cual se facilita el reconocimiento y el seguimiento de la motocicleta entre varios vehículos. Se utiliza en estos casos, la sustracción de la imagen del fondo para determinar los objetos en movimiento. Algunos trabajos utilizan la detección del casco para identificar con posterioridad las motos. (Chiverton, 2012).

La mayoría de estos trabajos, se soportan en la extracción de características propias de las motocicletas, tales como elementos circulares, HOG (histograma de gradientes) para el color o la forma, la relación de aspecto y la estimación de las áreas dentro de regiones de interés. Dicha literatura abarca desde 1985 hasta el año 2016, donde se utilizan Redes Neuronales Convolucionales para la detección de ellas en situaciones de tráfico denso. (Huynh & Hamamoto, 2016).

En este momento, la precisión máxima que se encuentra en el estado del arte para la detección y clasificación de motocicletas es del 97,8%.

b. Detección y Reconocimiento del uso del casco de seguridad en Motociclistas.

El trabajo más importante en esta área fue desarrollado por Chiverton (Chiverton, 2012). Este autor propone la siguiente metodología:

- Detección de un objeto en movimiento
- Aislamiento de la región de interés. Cálculo de su histograma después de la aplicación de varios operadores.
- Determinación del uso del casco por medio del análisis de características de la imagen transformada. En especial hacen uso del brillo que genera el material del casco para determinar su presencia en la imagen. Estas características son utilizadas por un clasificador SVM con un kernel Lineal para determinar la presencia de él.
- Seguimiento de la motocicleta haciendo uso de filtros de Kalman.

Chiverton, utiliza la sustracción de la imagen del fondo para la detección del objeto en movimiento. Se utiliza con posterioridad, la relación del aspecto (ancho contra largo) para una región de interés utilizando un rectángulo de encerramiento en la imagen. Dicha relación permite clasificar a un objeto como moto o de otra clase.

Se procede a aislar la región de interés donde se encuentra la cabeza del motociclista y se utiliza un proceso de extracción de características para determinar el uso del casco o la ausencia de este. Entre esas, el histograma del nivel de grises calculado sobre la zona de interés, el histograma del nivel de grises sobre cuatro cuadrantes en los cuáles se divide la región donde potencialmente se puede ubicar el casco, con lo cual se construye un vector. Así mismo, se construye un histograma de gradientes orientados (HOG) con una máscara de diferencia finita y otro HOG utilizando los gradientes máximos de color (i.e la reflectividad del casco es útil en esta tarea).

Finalmente, el autor de este estudio decide implementar un algoritmo de seguimiento de las motos durante los diversos frames que contiene el video. Para tal fin, hace uso de filtros Kalman para establecer la posición relativa de los píxeles de la imagen original a través del tiempo. La precisión para la detección de motocicletas, es del 83%, mientras que para la detección del casco es del 85%.

En otra investigación, Waranusast (Waranusast, Bundon, & Pattanathaburt, 2013) emplea una metodología similar a la utilizada por Chiverton.

Ese mismo autor en (Waranusast, Timtong, Bundon, & Tangnoi, 2014) utiliza una variación metodológica para detectar pasajeros. Hecho esto, separa la región de la cabeza para cada pasajero y se construye una región dividida en cuadrantes, tomando como centroide el punto medio de la misma. Se utiliza con posterioridad sobre esta región la extracción de características propias del uso del casco, tales como las circularidades de arco en los cuadrantes uno, dos y cuatro. Se puede observar dicha técnica a continuación:

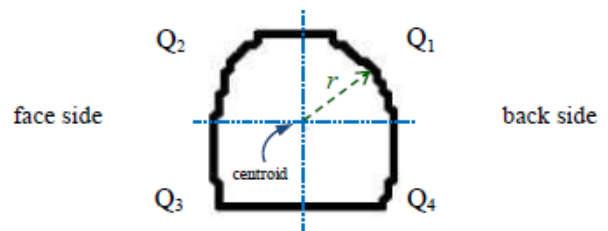


Ilustración 1-Cuadrantes en los cuáles se divide la región de la cabeza del motociclista.

Posteriormente se calculan los histogramas de las intensidades medias de grises para cada uno de los cuadrantes. Aquellas imágenes con casco, presentan mayores diferencias entre las intensidades en el cuadrante Q3 y Q2. No es así,

cuando no hay uso de él. Finalmente se calcula la diferencia de intensidades en la escala de color en el cuadrante Q3. Cuando se hace uso del casco, la variación en este cuadrante es sustancial, mientras que no es significativa en omisión de este.

Dicha investigación utiliza K-Nearest Neighbors (KNN) como algoritmo clasificador y se obtiene una precisión del 83,82% para la detección y Reconocimiento del uso del casco de seguridad en Motociclistas. Dicho valor aumenta hasta el 89% cuando la región de la cabeza se “corta” manualmente y se alimenta el modelo entrenado con un dataset de prueba.

Silva (Silva, y otros, 2013) utiliza para la detección y clasificación de las motos un método idéntico a los trabajos anteriores. Esto es, la sustracción de la imagen del fondo con su posterior segmentación. Se hacen uso de algunos filtros durante el preprocesamiento de la imagen para mejorar su calidad. Se puede observar en la siguiente imagen, la metodología adoptada por estos autores:

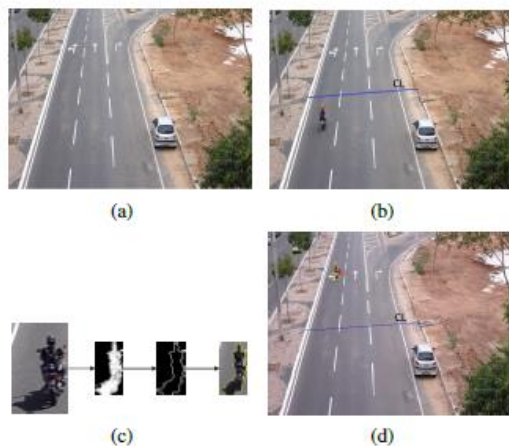


Ilustración 2-(a) Imagen de Fondo (b) Loop Virtual para detección del objeto (c) Sustracción de (a) y posterior segmentación (d) Resultado final con su correspondiente conteo (Silva, y otros, 2013).

Los autores proceden a la detección del casco, circunscribiendo a una región de interés (ROI) la extracción de características propias de este. Entre esas, hacen uso de la transformada circular de Hough para la detección de objetos circulares, así como de HOG relacionados con los gradientes de color y forma del casco. Finalmente se utiliza un descriptor híbrido, en la cual se umbraliza la imagen con OTSU y después se hace uso de Sobel para la detección de los bordes. Dicho descriptor mejora la aplicación de la transformada de Hough y permite obtener mejores resultados. Se utilizan como algoritmos clasificadores Bayes Nativo, Bosques

Aleatorios y SVM. La precisión de este estudio se maximizó al utilizar los bosques con un valor de 94,23% para la clasificación correcta en el uso del casco y disminuyó hasta el 89,13% con SVM.

En otro paper escrito por Silva (Silva, Aires, & Veras, Helmet detection on Motorcyclists using image descriptors and Classifiers, 2014) utilizan una metodología análoga a la ya presentada en este literal. Después de sustraer la imagen de fondo, se segmenta la imagen del objeto y se utiliza un vector de características para clasificarlo entre motocicleta o no motocicleta. Se procede a la detección del casco, determinando una región de interés, de la cual se extraen características propias de la imagen y se procede con su clasificación. Entre las novedades que se introducen en este estudio frente a los anteriores es el uso de una función de Transformada de Ondulación (WT). Está es una variación de la transformada de Fourier para el procesamiento de imágenes. Sobre cualquier imagen digital es posible computar los coeficientes de la Transformada de Ondulación. Tal transformada permite la extracción de características propias de personas que hacen uso del casco y que con posterioridad alimentarán clasificadores de inteligencia artificial. Se puede visualizar a continuación un ejemplo de la aplicación de la Transformada de Ondulación (WT).

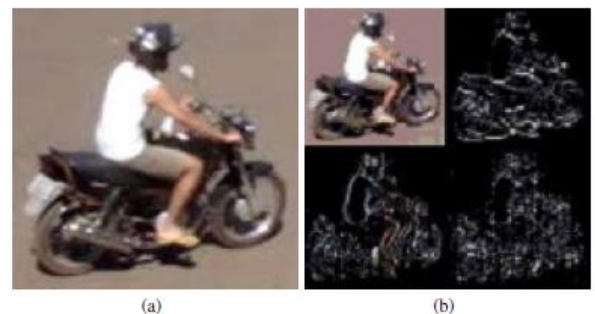


Ilustración 3- (a) Imagen Original (b) Resultado de la aplicación de la Transformada de Ondulación (WT)

En este estudio, Silva hace uso de clasificadores SVM, Perceptrones Multicapa (MLP), Funciones de Red radiales (RBFN), Bayes Nativo, Bosques Aleatorios y KNN. Para este caso, la mejor precisión la obtiene con el uso de MLP con un valor en la clasificación correcta en el uso del casco del 91,37%.

4. Metodología Propuesta

La metodología propuesta se puede describir en el siguiente marco:

- I. Detección de objetos en el flujo vehicular.
- II. Clasificación binaria para identificación de las motocicletas.
- III. Ubicación y Extracción de la moto desde la imagen original.
- IV. Extracción del ROI (región de interés) del motociclista y construcción del contorno de la zona de la cabeza del conductor, así como de la coordenada H en el stack HSB (hue, saturation and brightness) de la región de interés.
- V. Clasificación binaria de estas imágenes para identificación del uso del casco.

Se debe aclarar que el procesamiento de las imágenes, así como los modelos de inteligencia artificial que se utilizaron durante esta investigación se corrieron en un portátil común (Core i5-4210U @1,70 GHZ-12GB en RAM).

El marco metodológico, así como el problema a resolver, usualmente determinan la tipología de la captura de la imagen que se empleará para la investigación correspondiente. En este caso, se diseñó una captura fácilmente replicable con un bajo costo, con lo que se podrá en el futuro cercano, incrementar sustancialmente los puntos de captura para abarcar una mayor proporción de la población a estudiar. En ese sentido, las imágenes se capturaron de forma frontal y superior al tráfico vehicular en la ciudad de Bogotá, abarcando la totalidad de los carriles de la calzada. Las imágenes se tomaron en un puente peatonal, a cuatro metros con diez centímetros (4,1 mts) de altura sobre la rasante de la vía, entre las 12 pm y 3 pm, en un día nublado. Se utilizó una cámara GOPRO 4 de 12 MP sobre un trípode con una inclinación de -40° sobre el eje horizontal. Las fotos se tomaron en forma de ráfaga, de tal forma que se hizo una captura por segundo. El dataset original estuvo compuesto por 10.120 imágenes, de las que solo 820 tuvieron objetos en ella. Si bien, la forma de captura propuesta disminuye las posibilidades de encontrar oclusión en los objetos que se presentan en ella, también supone un reto mayor, para la identificación apropiada de los objetos, en la medida en que puede existir una gran cantidad de ellos en un momento dado (i.e congestión vehicular).

Se exhibe a continuación, la geometría de la imagen aplicada en este estudio.



Ilustración 4-Ejemplo de una imagen capturada según la propuesta

I. Detección de objetos en el flujo vehicular.

Para alcanzar satisfactoriamente la detección de objetos en el flujo vehicular, se sustrae el fondo de la imagen original, buscando detectar los objetos en movimiento en ella. Con posterioridad, se selecciona una Región de Interés donde la imagen en RGB se convierte a escala de grises (8-bits). Con posterioridad, se mejora el contraste de la foto, con el fin de reducir la incidencia de la luz sobre los objetos de la imagen.

Hecho esto, la imagen se umbraliza utilizando la técnica de momentos propuesto por Tsai. (Tsai, 1985). La umbralización se lleva cabo con el fondo invertido. Dicha técnica sumada a la sustracción del fondo, aumenta el principio de preservación del momento para los objetos en la foto. El resultado de este procesamiento se observa a continuación:

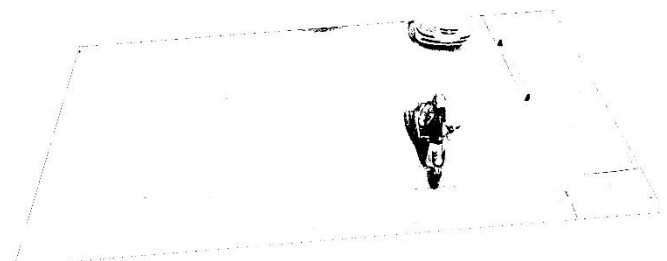


Ilustración 5-Imagen Resultante después de sustraer el fondo, seleccionar el ROI, mejorar el contraste, y umbralizar con Momentos

A continuación, se analizan las partículas que se encuentran en la imagen, para lo cual se segmenta, y se realzan los bordes de los objetos que aparezcan en ella. Para efectos de buscar las partículas de interés, se determinó que el tamaño que ocupa la moto en la región de interés (ROI) varía entre el 5% al 10% de la misma, con lo cual, objetos cuyo tamaño sea inferior, serán descartados por el análisis de partículas. El algoritmo, construye la morfología del objeto automáticamente por medio de la detección del borde. Dicho procesamiento, arroja una imagen como la que se exhibe a continuación:



Ilustración 6-Análisis de Partículas sobre la imagen umbralizada

Las técnicas descritas con anterioridad permiten identificar y detectar los objetos significativos en el flujo vehicular para la geometría de la captura de imagen propuesta. Para el procesamiento de estas imágenes se utilizaron las librerías de ImageJ sobre Python.

II. Clasificación binaria para identificación de las motocicletas.

La morfología de los objetos que se encuentran en las imágenes procesadas, permite la aplicación de modelos de inteligencia artificial para clasificar entre unos y otros. En especial, después de aplicar las técnicas de procesamiento imágenes descritas con anterioridad, la morfología permite distinguir entre aquellos objetos que son motos y aquellos que no lo son. Dicha clasificación se construiría de mejor forma con una red neuronal convolucional que utilice aprendizaje profundo para construir las capas de la red con las características morfológicas de las propias imágenes. Tal aproximación ya ha sido intentada por diversos autores y es a la fecha, el clasificador más preciso de imágenes (Krizhevsky, Sutskever, & Hinton, 2012). No obstante y debido las limitaciones que impone el tamaño del dataset, no es posible construir

una red convolucional propia. De tal forma que se hace uso de tres modelos pre-entrenados de aprendizaje profundo distintos para extraer las características de las imágenes procesadas.

Dichos modelos corresponden a iter_45, (Turi, 2014) Inception-V3 Network (Szegedy, 2015) y Full ImageNet Network (Ioffe & Szegedy, 2015). Cada uno de ellos ha sido entrenado con millones de imágenes en RGB y Grises y cuyas capas permiten la identificación de la morfología, el cambio de color, la detección de bordes y en general, se utilizan para el reconocimiento de imágenes multiclase a gran escala. La aplicación de cada uno estos modelos a las imágenes del Dataset, permite la extracción de las características (i.e features) propias de cada una ellas. Las características se representan por un vector.

En consecuencia, el Dataset de Prueba incorpora entonces la imagen original, la imagen con una transformación geométrica que requiere el paquete de Software de inteligencia artificial, su ID, tres columnas de características propias (i.e cada característica es un vector que se obtiene de aplicar cada modelo pre entrenado de Deep Learning sobre ella) y la etiqueta. El Dataset de prueba para utilizar en los distintos modelos está compuesto por 109 imágenes. La mitad para la clase de moto y la otra para la clase no moto, las cuales fueron etiquetadas con anterioridad. El dataset de prueba está compuesto por 35 imágenes con la misma estructura. Algunas imágenes que componen el Dataset de entrenamiento, se pueden observar a continuación:

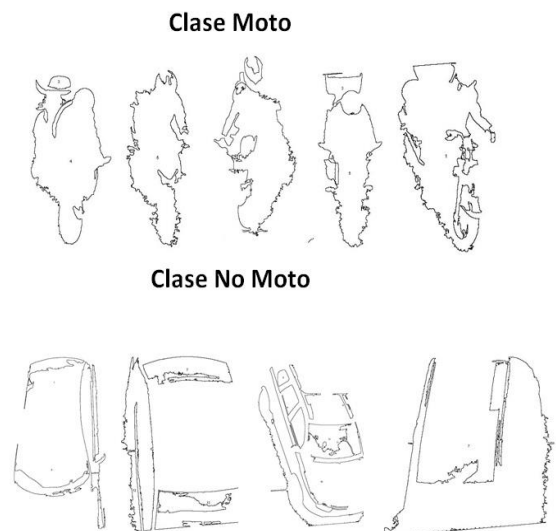


Ilustración 7-Ejemplos de Imágenes que componen el dataset de entrenamiento

Con los vectores de características y la etiqueta como objetivo, se construyen tres clasificadores distintos. Un clasificador con arboles de decisiones, otro con Support Vector Machines (SVM) y el último con Boosted Trees. Para esto, se hace uso de IPython sobre Jupyter con librerías de Graphlab. Los clasificadores se evalúan en función de su grado de precisión ("accuracy"), su área bajo la curva ("auc") y el puntaje f1 ("f1_score").

Para el primero, la precisión es del 94,28%, el "auc" es de 0,9375 y el puntaje f1 es de 0,933. Para el segundo (SVM), la precisión es del 91,15% con auc del 0,9487 y puntaje f1 de 0,937.

Finalmente, el último alcanza una precisión del 97,15% con "auc" del 0,9687 y puntaje f1 de 0,967. Estos son los mejores resultados alcanzados por cualquiera de los tres clasificadores utilizados y se debe en principio, a que el modelo de Boosted Trees se basa en una colección base de sistemas de aprendizaje (i.e clasificadores basados en árboles de decisión) que pueden modelar relaciones no lineales entre las características y el objetivo (i.e vectores de características contra la etiqueta de la clase). (Friedman, Hastie, & Tibshirani, 2000)

III. Ubicación y Extracción de la moto desde la imagen original.

Una vez se etiqueta el objeto de la imagen como motocicleta, se procede por medio de un script a duplicar el espacio donde se encuentra la misma en la imagen preprocesada (i.e antes de ser convertida a escala de grises) y se suma el fondo nuevamente, para obtener el objeto en RGB. Se configuran dos Datasets. Uno con imágenes con el fondo sustraído y otro que contiene las imágenes en RGB. El procesamiento de estas imágenes se hace utilizando librerías de ImageJ sobre Python.

IV. Extracción del ROI (región de interés) del motociclista y construcción del contorno de la zona de la cabeza del conductor, así como de la coordenada H en el stack HSB (hue, saturation and brightness) de la región de interés.

En la medida en que la geometría de la captura de la imagen propuesta, supone que una motocicleta puede tener uno o dos pasajeros y

que la última circunstancia puede generar la oclusión de un casco sobre otro, se determinó que para esta primera fase de la investigación, el análisis del procesamiento de imágenes propuesto solo se aplicará a aquellas motos donde exista únicamente una persona en ella. Dicha selección en el dataset se hizo de forma manual.

Con la imagen original en RGB y con el fondo sustraído de la motocicleta, se procede a partir la misma por tercios y se duplica el primer tercio, siendo esta la región de interés (ROI) del motociclista, ya que allí es donde se debe encontrar el casco de seguridad. Se muestra a continuación una imagen tipo para el proceso descrito con anterioridad:



Ilustración 8-Extracción del ROI en RGB



Ilustración 9- Extracción del ROI con el fondo sustraído

Con el fin de identificar la geometría del casco, se convierte a escala de grises las imágenes con el fondo sustraído. Con posterioridad, la imagen se umbraliza utilizando la técnica de máxima entropía propuesta por Kapur. (Kapur, Sahoo, & Wong, 1985). Dicha técnica garantiza la máxima entropía intraclases, con lo cual se acentúa la forma del casco frente a los otros objetos en la imagen. Nuevamente, la umbralización se lleva cabo con el fondo invertido y apilando el histograma. Se hace con posterioridad un llenado morfológico con su posterior cierre para asegurar de forma óptima

la consecución de la forma del casco. Se muestra a continuación la imagen resultante de esta operación:

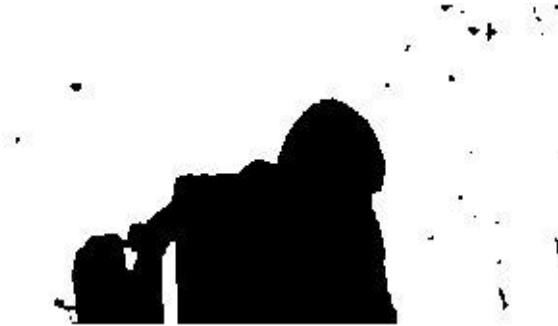


Ilustración 10- Imagen Umbralizada con la técnica de máxima entropía después de aplicar el llenado y cierre morfológico

Finalmente se aplica el operador de Sobel, para resaltar los cambios drásticos de intensidad en la imagen y como consecuencia, obtener los bordes de la figura. Dicho proceso se exhibe a continuación:

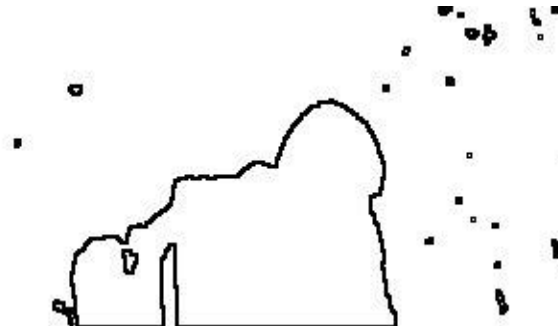


Ilustración 11-Aplicación del Operador de Sobel para la detección de bordes

Con los píxeles que determinan el contorno, se corta la imagen en el dataset en RGB y se vacía todo el contenido que esté afuera de ella. Con posterioridad se calcula la coordenada Hue (cambio de color) en el stack HSB. Dicho procedimiento se muestra a continuación:

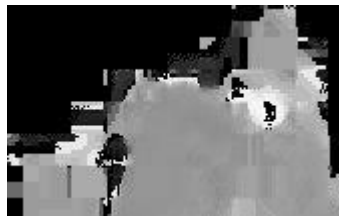


Ilustración 12-Extracción de la imagen del Dataset de RGB después de calcular la coordenada Hue

Dentro de la totalidad del dataset que se recolectó, no se consiguió ninguna imagen de

un motociclista que no estuviera portando casco, con lo cuál, fue necesario conseguir un dataset alternativo. Dichas imágenes se tomaron en la misma ciudad, en distinto sitio, con la misma cámara y en la misma configuración. No obstante, se fotografiaron ciclistas, ya que la inobservancia de la norma en este grupo poblacional en la ciudad de Bogotá, es presumiblemente más alto que el de los motociclistas. La extracción de la región de interés, para cada ciclista se hizo de forma manual y se aplicaron las mismas técnicas de procesamiento de imágenes descritas en este numeral. Se muestra a continuación la aplicación de las técnicas de procesamiento de imágenes sobre el dataset alternativo:



Ilustración 13-Imagen Original para Dataset Alternativo



Ilustración 14- Extracción del ROI del ciclista



Ilustración 15-Extracción del ROI con el fondo sustraído



Ilustración 16-Aplicación de las Técnicas de procesamiento de Imágenes (i.e Umbralización, Cierre morfológico y Sobel)



Ilustración 17- Cálculo de la coordenada Hue

Se debe mencionar que la morfología no es buen clasificador del uso del casco según la geometría de captura de la imagen utilizada en esta investigación. Para el caso de la identificación del casco, la morfología es distinta para cada objeto en la vía según la posición en la que se encuentre este en la imagen original frente al dispositivo de captura (i.e si la imagen se captura de forma frontal a la cámara, no se generará la forma circular propia del elemento de seguridad) con lo cual, este descriptor resulta casi inservible para determinar el uso del casco. Se combina a utilizar este descriptor (morfología), conjunto con el Histograma de Gradientes (HOG) o el cálculo de momentos de Hu, para imágenes con geometría de captura de la imagen similar a las expuestas en el trabajo de Chiverton (Chiverton, 2012).

Tal y como lo demuestran las imágenes expuestas, el cambio de la intensidad de color en la coordenada Hue es de mayor valor en aquellas personas que no utilizan casco. Esto es por el contraste que genera el color del pelo con el color de cara. Dicha variación es menor en aquellos que usan casco, ya que la visera del mismo, produce una variación insignificante con el color del casco. Esta circunstancia se puede utilizar como un atributo clasificador binario para el dataset existente. Se debe mencionar que este algoritmo resulta casi inservible en personas que porten capuchas o gorros con ellos. Para esta circunstancia, se

aconseja modificar la geometría de la captura de la imagen.

V. Clasificación binaria de estas imágenes para identificación del uso del casco.

Nuevamente, la aplicación de distintas técnicas de procesamiento de imágenes simplifican el funcionamiento de los modelos pre-entrenados de aprendizaje profundo para extraer las características de los cambios de intensidad en las coordenadas Hue del Dataset Propuesto.

Se utilizan nuevamente los tres modelos, iter_45, Inception-V3 Network y Full ImageNet Network y se calculan tres vectores de características distintas para cada modelo.

El dataset de prueba está compuesto por 77 imágenes, 40 de ellas con casco y 37 de ellas sin él. Nuevamente, todos los datos son etiquetados con anterioridad y tienen la estructura del dataset que se utilizó para la creación del clasificador de motocicletas. El dataset de prueba está compuesto por 34 imágenes. Algunas imágenes que componen el dataset de prueba se pueden visualizar a continuación.

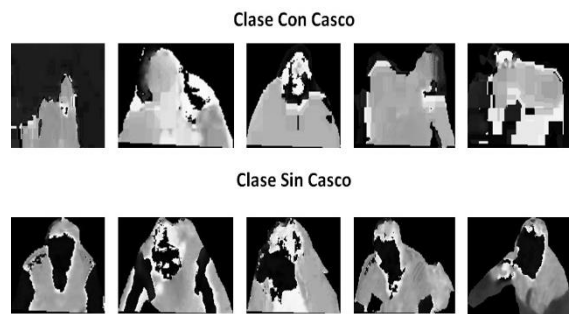


Ilustración 18-Ejemplo de imágenes que componen el dataset de entrenamiento

Considerando los resultados obtenidos para los modelos de inteligencia artificial que se utilizaron con ocasión del clasificador de motos, se utiliza exclusivamente el modelo de Boosted Trees. Se construye por tanto un clasificador binario para la identificación de la clase respectiva (uso del casco o no uso del casco).

Este clasificador se evalúa en función de su grado de precisión ("accuracy"), su área bajo la curva ("auc") y el puntaje f1 ("f1_score").

Entrenado y evaluado el modelo, arroja una precisión del 85,29%, el “auc” es de 0,8275 y el puntaje f1 es de 0,878.

5. Resultados

Se muestra a continuación los resultados del modelo clasificador binario para identificar motocicletas basado en Boosted Trees, así como su matriz de confusión:

```
{'accuracy': 0.9714285714285714, 'auc': 0.96875,
  target_label  int
  predicted_label int
  count         int

Rows: 3

Data:
+-----+-----+-----+
| target_label | predicted_label | count |
+-----+-----+-----+
| 1            | 0              | 1     |
| 1            | 1              | 15    |
| 0            | 0              | 19    |
+-----+-----+-----+
[3 rows x 3 columns], 'f1_score': 0.96774193548
```

Ilustración 19- Resultados del Clasificador de motos basado en Boosted Trees

Según la revisión presentada en (Zezhi, Ellis, & Velastin, 2011), la precisión para el mejor clasificador de motocicletas es de 97,9%. El clasificador propuesto, después de aplicar las técnicas de procesamiento de imágenes alcanza una precisión muy similar. No obstante, el propuesto en este trabajo incorpora dos ventajas importantes frente al que se enuncia. La primera de ellas es el tiempo de convergencia de los modelos de inteligencia artificial. Mientras que para la investigación a la cuál se hace alusión, se requirió de una gran capacidad de cómputo y una gran cantidad de tiempo para que el modelo de Deep learning entrenara por la gran cantidad de características que se pueden derivar de una imagen RGB con una amplia resolución, el propuesto en este trabajo lo pudo hacer en un tiempo inferior a cinco minutos en un portátil común.

De otro lado, la aplicación de modelos pre entrenados en la construcción del clasificador propuesto en esta investigación, supone una limitación en la precisión que puede alcanzar el mismo. Es posible que con un dataset lo suficientemente grande se pueda entrenar un modelo específico de Deep learning que permita mejorar la precisión en la detección de motos. Dicho modelo, puede ser utilizado

inclusive, en la clasificación posterior de objetos muy similares, tales como las bicicletas y las motos.

Finalmente se exhibe a continuación los resultados del modelo clasificador binario para identificar casco de seguridad basado en Boosted Trees, así como su matriz de confusión:

```
{'accuracy': 0.8529411764705882, 'confusion_matrix':
  target_label  int
  predicted_label int
  count         int

Rows: 4

Data:
+-----+-----+-----+
| target_label | predicted_label | count |
+-----+-----+-----+
| 1            | 1              | 18    |
| 1            | 0              | 2     |
| 0            | 1              | 3     |
| 0            | 0              | 11    |
+-----+-----+-----+
[4 rows x 3 columns], 'f1_score': 0.8780487804878048,
```

Ilustración 20- Resultados del Clasificador de casco de seguridad basado en Boosted Trees

Según la revisión presentada en (Silva, y otros, 2013), la precisión para el mejor clasificador del uso de casco de seguridad es 94,23%. El clasificador propuesto en esta investigación alcanza el 85,29%. Aunque la metodología propuesta por (Silva, y otros, 2013) resulta similar a la propuesta en esta investigación, en la referencia no se hace uso del Stack HSB, ni tampoco de algoritmos de Deep learning para extraer las características de la imagen.

Al igual que para el anterior clasificador, un dataset de mayor tamaño que permita construir un modelo de aprendizaje profundo propio para su uso en el clasificador, supondrá un incremento sustancial de la precisión para el algoritmo propuesto.

6. Conclusiones y Trabajo Futuro

La metodología presentada en este artículo permite la detección de motocicletas en el flujo vehicular. Así mismo, la geometría para la captura de la imagen que se utiliza proporciona una fácil replicación de este experimento en distintos sitios para ayudar a caracterizar la población de motociclistas en Colombia. Esto le permitirá a las autoridades correspondientes tomar las medidas de política pública

necesarias para mitigar y reducir los altos costos que produce la accidentalidad vial.

Las precisiones que se alcanzaron en los clasificadores tanto para la detección de motocicletas como para la detección del casco de seguridad resultan comparables con el estado del arte y suponen una configuración metodológica novedosa que se sustenta en partes iguales en las técnicas de procesamiento de imágenes, así como en los modelos de inteligencia artificial. Como trabajo futuro, se debe calibrar los modelos que aquí se presentan con un tamaño mayor de datos, así como también se debe resolver el problema de la correcta identificación de la cantidad de pasajeros en las motocicletas.

7. Referencias

- A. M., L. X., & T. B. (2013). On-Road Approaching Motorcycle Detection and Tracking Techniques: A Survey. *IEEE International Conference on Control System, Computing and Engineering*, (págs. 63-68). Penang.
- Blanc, C., Aufrere, L., Malaterre, J., & Alison, J. (2004). Obstacle Detection and Tracking by millimeter wave radar. *IEEE Symposium on Intelligent Autonomous Vehicle (IAV)*. Lisboa.
- Chiu, C.-C., Ku, M.-y., & Chen, H.-T. (2007). Motorcycle detection and Tracking System with Occlusion Segmentation. *8th International Workshop on Image Analysis for Multimedia Interactive Services* (págs. 21-27). Taiwan: IEEE.
- Chiverton, J. (2012). Helmet presence classification with motorcycle detection and Tracking. *IET Intelligent Transport Systems vol:6*, 259-269.
- Chung-Cheng, C., Min-Yu, K., & Hung-Tsung, C. (2007). Motorcycle Detection and Tracking System with Occlusion Segmentation. *Image Analysis for Multimedia Interactive Services* (págs. 32-32). WLAMIS'07.
- Dirección de Epidemiología y Demografía, Ministerio de Salud. (2015). *Mortalidad y lesiones por accidentes de transporte en Colombia, 2013-2014*. Bogota: Minsalud.
- El Espectador. (4 de Abril de 2015). *Colombia, un país que se transporta en moto*. Obtenido de El Espectador: <http://www.elespectador.com/noticias/economia/colombia-un-pais-se-transporta-moto-articulo-553177>
- Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression a statistical view of boosting. *The annals of statistics*, 337-407.
- Herbert, M. (2000). Active and Passive range sensing for robotics. *Robotics and Automation Proceedings 2000*, (págs. 102-110). Vancouver.
- Huynh, C., & Hamamoto, K. (2016). Convolutional Neural Network for Motorbike Detection in Dense Traffic. *Communications and Electronics (ICCE), 2016 IEEE Sixth International Conference on* (págs. 27-29). IEEE.
- Instituto Javeriano de Salud Pública. (2015). *Patrón de mortalidad en motociclistas en Colombia 2000-2014*. Bogota: Universidad Javeriana.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*. Obtenido de <https://github.com/dmlc/mxnet-model-gallery/blob/master/imagenet-21k-inception.md>
- Kanhere, N., Birchfield, S., Sarasua, W., & Khoemi, S. (2010). Traffic Monitoring of Motorcycles During Special Events Using Video Detection. *Transportation Research Board Journal of the Transportation Research Board vol 2160*, 69-76.
- Kapur, J., Sahoo, P., & Wong, A. (1985). "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram". *Graphical Models and Image Processing*, 273-285.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 25-37.

- Ling, B., Gibson, P., & Middleton, D. (2013). Motorcycle Detection and counting using stereo camera, IR camera and microphone array. *Sensor, Mesh and Ad Hoc Communications and Networks*, 86630.
- Loaiza, C. (30 de Mayo de 2012). *Motos: Accidentes fatales*. Obtenido de Sura Blog: <http://www.sura.com/blogs/autos/motos-accidentes-fatales.aspx>
- Meeselodi, S., Modena, M., & Cattoni, G. (2007). Vision-Based Bicycle/Motorcycle classification. *Pattern recognition Letters*, 1719-1726.
- Min-Yu, K., Chung-Cheng, C., Hung-Tsung, C., & Hong, S. (2008). Visual Motorcycle Detection and Tracking Algorithms. *WSEAS transactions on electronics*, 121-131.
- Nashashibi, F., & Bargeton, A. (2008). Laser-based vehicle tracking and classification using occlusion reasoning and confidence estimation. *Intelligent Vehicles Symposium*, (págs. 847-852).
- Nguyen, P., & Le, H. (2008). A Multi-modal particle filter based motorcycle Tracking System. *PRICA 2008: Trends in AI*, 818-829.
- Roess, R., Prassas, E., & McShane, W. (2004). *Traffic Engineering*. Washington: Pearson.
- Sasikala, G., Padol, K., Katekar, A., & Dhanasekaran, S. (2015). Safeguarding of Motorcycles Through Helmet Recognition. *International Conference on Smart Technologies and Management* (págs. 609-612). Chennai: IEEE.
- Silva, R., Aires, K., & Veras, R. (2014). Helmet detection on Motorcyclists using image descriptors and Classifiers. *27th SIBGRAPHI Conference on Graphics, Patterns and Images* (págs. 141-148). Picos: IEEE.
- Silva, R., Santos, T., Kalyf, A., Veras, R., Aires, K., & Soares, A. (2013). Automatic detection of motorcyclist without helmet. *Latin American Computing Conference*. Rio de Janeiro.
- Szegedy, C. (2015). Rethinking the Inception Architecture for Computer Vision. *arXiv preprint arXiv:1512.00567*. Obtenido de <https://github.com/dmlc/mxnet-model-gallery/blob/master/imagenet-1k-inception-v3.md>
- Tsai, W. (1985). "Moment-preserving thresholding: a new approach". *Computer Vision, Graphics, and Image Processing*, 377-393.
- Turi. (2014). *Modelo Iter_45*. Obtenido de http://s3.amazonaws.com/GraphLab-Datasets/deeplearning/imagenet_model_iter45
- Valbuena, S. (2012). *Muertes y lesiones no fatales por accidentes de transporte, Colombia, 2011*. Bogota: Instituto Nacional de Medicina Legal y Ciencias Forenses.
- Waranusast, R., Bundon, N., & Pattanathaburt, P. (2013). Machine Vision Techniques for Motorcycle Safety Helmet Detection. *International Conference on Image and Vision Computing*, (págs. 13-18). New Zealand.
- Waranusast, R., Tintong, V., Bundon, N., & Tangnoi, C. (2014). A computer Vision Approach for Detection and Counting of Motorcycles Riders in University Campus. *Proceedings of the IEEE Congress . Phitsanulok: IEEE*.
- Wen-Chung, C., & Kuo-Jung, H. (2010). Vision-based side vehicle detection from a moving vehicle. *ICSSE 2010, International Conference*, (págs. 553-558).
- Zezhi, C., Ellis, T., & Velastin, A. (2011). Vehicle Type categorization: A comparison of classification schemes. *ITSC 2011 14th International IEEE Conference*, (págs. 74-79).

APÉNDICE

A. Archivos

Los scripts hechos en python, los datasets, así como los modelos entrenados de inteligencia artificial se suben al siguiente repositorio:

https://github.com/mavaron/Tesis_MSc

B. Lista de Imágenes

Ilustración 1-Cuadrantes en los cuáles se divide la región de la cabeza del motociclista.3

Ilustración 2-(a) Imagen de Fondo (b) Loop Virtual para detección del objeto (c) Sustracción de (a) y posterior segmentación (d) Resultado final con su correspondiente conteo (Silva, y otros, 2013).4

Ilustración 3- (a) Imagen Original (b) Resultado de la aplicación de la Transformada de Ondulación (WT) ...	4
Ilustración 4-Ejemplo de una imagen capturada según la propuesta	5
Ilustración 5-Imagen Resultante después de sustraer el fondo, seleccionar el ROI, mejorar el contraste, y umbralizar con Momentos	5
Ilustración 6-Análisis de Partículas sobre la imagen umbralizada	6
Ilustración 7-Ejemplos de Imágenes que componen el dataset de entrenamiento	6
Ilustración 8-Extracción del ROI en RGB	7
Ilustración 9- Extracción del ROI con el fondo sustraído	7
Ilustración 10- Imagen Umbralizada con la técnica de máxima entropía después de aplicar el llenado y cierre morfológico	8
Ilustración 11-Aplicación del Operador de Sobel para la detección de bordes.....	8
Ilustración 12-Extracción de la imagen del Dataset de RGB después de calcular la coordenada Hue	8
Ilustración 13-Imagen Original para Dataset Alternativo	8
Ilustración 14- Extracción del ROI del ciclista	8
Ilustración 15-Extracción del ROI con el fondo sustraído	8
Ilustración 16-Aplicación de las Técnicas de procesamiento de Imágenes (i.e Umbralización, Cierre morfológico y Sobel)	9
Ilustración 17- Calculo de la coordenada Hue	9
Ilustración 18-Ejemplo de imágenes que componen el dataset de entrenamiento	9
Ilustración 19- Resultados del Clasificador de motos basado en Boosted Trees.....	10
Ilustración 20- Resultados del Clasificador de casco de seguridad basado en Boosted Trees.....	10