



UNIVERSIDAD NACIONAL DEL LITORAL

Facultad de Ingeniería y Ciencias Hídricas

PROYECTO FINAL DE CARRERA

INGENIERÍA INFORMÁTICA

Informe Final

**Desarrollo de un sistema de detección de pérdidas de atención
en la conducción basado en visión computacional**

Alumno: Farré, Gonzalo

Director: Martínez, César

Co-Director: Albornoz, Enrique Marcelo

Santa Fe, Abril 2020

Agradecimientos

Agradezco particularmente a mi familia y a mis compañeros del colegio secundario que siempre me dieron su apoyo y estuvieron conmigo hasta el final. A mi tía Silvia y a mi abuela Elena que en particular siempre estuvieron alentándome y demostrándome su interés por mi desempeño.

A mis profesores, Albornoz Enrique Marcelo y Martínez César, los cuales aprobaron desde el día uno mi proyecto, y me acompañaron con ímpetu desde mi regreso del Intercambio Académico, hasta el momento.

A la Universidad de Playa Ancha de Valparaíso, por hacerme vivir una experiencia totalmente nueva, con personas increíbles las cuales marcaron mi vida de forma tanto académica como personal.

Dedicado también a la familia de Diveria S.R.L, que como bien menciono, más que una empresa es una familia, dedicada al desarrollo de software de calidad, la cual me acompañó a lo largo de mi primera experiencia laboral y lo que fue mi Práctica Profesional Supervisada.

Finalmente a mis compañeros y futuros colegas, Castello Facundo, Castillo Cesar, Deangeli Damian, Fagioli Gian, Grosso Federico, Lopez David, Loyarte Facundo, Salmeron Facundo, Senger Agustín, Sikh Lautaro, Tesolin Guido, Walczak Tomás y Yackel Francisco. A cada uno de ellos está dedicado este proyecto final de carrera ya que estuvieron en cada parcial, en cada nota, en cada clase, en cada festejo. Por siempre, gracias.

Índice general

Índice de figuras	8
1. Introducción	11
1.1 Justificación.....	11
1.2 Objetivos	13
1.2.1 Objetivo general	13
1.2.2 Objetivos específicos	13
1.3 Alcance y limitaciones	14
2. Diseño de la base de datos.	15
2.1 Posición de la cámara y especificaciones.	15
2.1.1 Posición.....	15
2.1.2 Especificaciones	18
2.2 Características generales de los videos.....	19
2.2.1 Datos técnicos.....	19
2.2.2 Características generales.....	20
2.2.3 Metodologías para el registro de videos.....	26
2.3 Otras consideraciones sobre la base de datos	26
3. Desarrollo del sistema	28
3.1 Sistema operativo y librerías	28
3.2 Diseño general del método	28
3.2.1 Detección de Rostro.....	29
3.2.2 Estimación de postura inicial.....	33
3.2.3 Evaluación de imágenes día/noche	35
3.2.4 Corrección Gamma	37
4. Evaluación de distracciones	40

4.1	Distracciones con pérdida de visión total	41
4.1.1	Buffer de distracciones.....	42
4.2	Distracciones con pérdida de visión parcial	42
4.3	Parpadeos y somnolencia	43
4.3.1	Índice de somnolencia.....	45
4.4	Diseño de alarmas temporales.....	46
4.4.1	Modelado de las distracciones	46
4.4.2	Desarrollo de alarmas progresivas.....	48
4.5	Resultados y discusión.....	50
4.5.1	Resultados sobre el conjunto inicial	50
4.5.2	Resultados sobre conjunto de prueba.....	53
5.	Interacción con el usuario y componentes visuales.....	57
5.1	Menú principal y videos de entrada.....	57
5.1.1	Interacción con usuario	57
5.1.2	Selección de video de entrada	58
5.2	Elementos visuales adicionales	59
5.3	Elaboración de informes.....	61
5.3.1	Creación de archivos.....	61
5.3.2	Creación de carpetas	62
5.4	Discusión.....	63
6.	Conclusiones y desarrollos futuros	64
	Referencias	67
	Plan del proyecto y manual de usuario.....	69

Índice de figuras

Fig. 2.1: Marcadores faciales brindados por Dlib.....	16
Fig.2.2: Conductores del primer dataset.....	16
Fig. 2.3: Conductores del segundo dataset.	17
Fig. 2.4: Conductores del tercer dataset.....	17
Fig. 2.5: Distintas posiciones de cámara	18
Fig. 2.6: Posición de cámara desde diferentes ángulos.	19
Fig. 2.7: Duración de los videos y distracciones en los mismos	20
Fig. 2.8: Contenido del archivo de texto que acompaña a cada video	24
Fig. 2.9: Ejemplos de pérdida de atención total sin percepción del camino	24
Fig. 2.10: Ejemplos de pérdidas de la atención con vista parcial del camino	24
Fig. 2.11: Capturas de distracciones leves para mujeres.....	26
Fig. 2.12: Capturas de distracciones Moderadas para varones.....	26
Fig. 2.13: Capturas de distracciones peligrosas para ambos sexos.....	27
Fig. 2.14: Capturas de distracciones en videos de prueba.....	27
Fig. 3.1: Diagrama de bloques del sistema.....	29
Fig. 3.2: Método de detección de rostros por clasificadores en cascada	31
Fig. 3.3: Ejemplo de vecindad en el método de Patrones Binarios Locales	31
Fig. 3.4: Marcadores faciales obtenidos mediante la biblioteca Dlib.	32
Tabla 3.1: Resultados de comparación entre Haar y LBP.....	33
Fig. 3.5: Mapeo del conjunto de 14 puntos 2D a un modelo 3D	34
Fig. 3.6: Ejemplo de Detección utilizando ángulos de Euler.....	35
Fig. 3.7: Ejemplos de Histogramas RGB	36

Fig. 3.8: Ejemplos de ROI tanto para una imagen de día y para una nocturna.	37
Fig. 3.9: Diferentes valores de Gamma junto con su curva logarítmica.....	38
Fig. 3.10: Capturas nocturnas con Correccion Gamma.....	39
Fig. 4.1: Diferencias entre distracciones Parciales y Totales.....	41
Fig. 4.2: Enumeración de puntos en Dlib.....	44
Fig. 4.3: Imagen simplificada de la relación entre clases	47
Fig. 4.4: Imagen de disparo de la alarma 1 en detección de distracciones leves.	48
Fig.4.5: Pantalla final con información de relevancia.....	49
Fig.4.6: Ejemplo de informe final de cada video	53
Fig. 4.7: Capturas pertenecientes a los videos de prueba.....	56
Fig. 5.1: Pantalla de bienvenida.	58
Fig. 5.2: Respuesta cuando el usuario ingresa una letra en lugar de un número.....	58
Fig.5.3: Listado de videos preconfigurados.	59
Fig. 5.4: Señalización visual de distracciones.	60
Fig. 5.5: Consulta al usuario sobre la creación del informe final.....	62
Fig. 5.6: Notificaciones sobre las carpetas creadas.	63

- Capítulo I -

1. Introducción

Se podría decir que se está viviendo en un mundo en constante evolución, y mientras la tecnología está acortando distancias continentales a través de un click, las distancias físicas siguen siendo las mismas. La sociedad se moviliza a través de vehículos motorizados cada vez más modernos día tras día, y a pesar de que la tecnología avance constantemente, la seguridad vial sigue siendo una de las principales causas de preocupación en lo que respecta a nuestra vida cotidiana.

Las distracciones al conducir son una de las principales causas de accidentes de los últimos tiempos, y a pesar de que éstos hechos se repitan una y otra vez, pocas son las cosas que se llevan a cabo para prevenirlos.

Es así que en el presente proyecto se planteará una alternativa para la detección y prevención de pérdidas de la atención al conducir mediante el análisis de distintas simulaciones de situaciones viales con el fin de relevar datos y concientizar al conductor a la hora de conducir.

1.1 Justificación

La utilización de automóviles para el traslado es un hecho totalmente habitual en la actualidad. Sin embargo, a pesar de los avances tecnológicos que pudieron llegar a darse en el campo automovilístico, la seguridad vial sigue siendo una de las principales preocupaciones que surge a partir de la gran cantidad de siniestros viales.

Según un estudio realizado por el Foro Internacional del Transporte [1], Argentina encabeza una lista de 38 países en términos de muertes por causas viales, donde 5000 fallecimientos son los que se registran anualmente a una tasa de 12,3 muertes por cada 100.000 habitantes. La tasa más pequeña la tienen países como Suecia, con una tasa de 2,7 defunciones por año cada 100.000 habitantes [2].

A partir de esto, es importante destacar una serie de datos que serán de utilidad para entender un poco más el contexto en el que se encuentra el proyecto y analizar la problemática que se nos presenta. De acuerdo con las estadísticas de CESVI (Centro de Experimentación, Investigación y Seguridad Vial) sobre la base del análisis de 4000 accidentes de tránsito ocurridos en los últimos 11 años, la mitad de los accidentes

fueron protagonizados por autos. Lo siguen los camiones (23%), pick-ups (9%), utilitarios y 4x4 (8%), micros (3%), motocicletas (3%) y otros (4%) [3].

A su vez, es importante destacar las principales causas de dichos accidentes, siendo 90% de los choques debidos al factor humano, 5% al vehículo y 5% consecuencia de las malas condiciones de caminos y el clima. Ahora, haciendo foco en la principal causa de accidentes, se mencionan las principales falencias humanas que se registraron en dichos escenarios, donde según la institución mencionada, 41% de las fallas humanas que llevan a los accidentes es la invasión de carril. Siguen: la distracción (19%), la velocidad inadecuada (16%), las maniobras abruptas (10%), distancia inadecuada (6%), prioridad de paso no respetada (4%) y cansancio (4%). Obviamente, estarían incluidas dentro de estas fallas las causadas por la ingesta de alcohol.

Las distracciones, razones principales para la pérdida de atención, son definidas según la Real Academia Española, como “cosa que atrae la atención apartándola de aquello a lo que está aplicada...”. Con esta definición, se puede catalogar una serie de distracciones comunes que eventualmente participan en los accidentes mencionados anteriormente, entre ellas se mencionan: utilización del teléfono celular, uso del cigarrillo, maquillaje, búsqueda de objetos, etc. [4].

Para ver como se le hace frente a éstos accidentes en la actualidad, juega un papel importante Bosch, empresa alemana vinculada a la construcción y distribución de sistemas para la industria automovilística, la cual ha desarrollado un sistema de detección de fatiga aplicable a casi cualquier tipo de auto, el cual se encarga de utilizar la información de distintos sistemas integrados del automóvil con el fin de evaluar las acciones del conductor y buscar signos de cansancio y sueño [5].

Lo que se pretende con este proyecto es, teniendo en cuenta las estadísticas mencionadas, construir un sistema que no sólo se encargue de las fatigas del conductor, sino que se ocupe de las distracciones del mismo (siendo las distracciones la causa del 19% de las falencias humanas contra sólo un 4% que aporta el cansancio y la fatiga). Se atacarán las distracciones mencionadas midiendo la posición de la cabeza del conductor y el tiempo que ésta se desvía de la conducción del automóvil. Detectando además, la somnolencia de la cual se ocupa Bosch.

El sistema irá evaluando, a través de una base de datos compuesta por videos, la dinámica temporal de la posición de la cabeza del conductor brindando diferentes tipos de "Alarmas" o puntos de control de acuerdo a la cantidad de tiempo que éste pierde su atención al conducir. Una vez finalizado cada video se mostrará un análisis con las distracciones observadas.

Con la realización de éste proyecto, se pretende de alguna manera crear conciencia de la existencia estos accidentes, mediante la evaluación de los videos planteados de modo que el conductor esté consciente de las conductas peligrosas al volante y trate de evitarlas en el futuro. También está en el campo de interés que la interfaz sirva a entidades gubernamentales para la enseñanza de la buena conducción y prevención de accidentes, como también que organismos tales como escuelas o institutos utilicen el sistema desarrollado con el fin de inculcar hábitos de la correcta conducción vial en las personas desde una edad temprana. Éste se podría decir que será el impacto social que podría provocar el proyecto, junto con un movimiento de concientización por parte de toda la comunidad de conductores.

1.2 Objetivos

1.2.1 Objetivo general

Desarrollar un sistema basado en visión computacional para detección de pérdida de atención en la conducción, mediante el análisis de la posición y orientación de la cabeza del conductor.

1.2.2 Objetivos específicos

- Relevar los requerimientos del sistema.
- Estudiar las diferentes propuestas reportadas en el estado del arte para la tarea abordada.
- Construir una base de datos adecuada para la tarea.
- Analizar y categorizar los diferentes tipos de distracciones que pueden encontrarse.
- Proponer un sistema de alarmas que indique la severidad de las distracciones.
- Diseñar y e implementar el sistema para la detección de pérdida de atención en la conducción.
- Evaluar el desempeño del sistema empleando un subconjunto de videos en la base de datos.
- Realizar informes periódicos con el fin de ir documentando el avance.
- Redactar un documento que integre todas las etapas del proyecto.

1.3 Alcance y limitaciones

El proyecto pretende el análisis de una base de datos de video con una cámara situada en un lugar estratégico del vehículo. Así, se puede decir que podría ser aplicado a todo vehículo de 2 ejes o más, sea auto, colectivo, camión, etc.

Además, se supondrá el mismo formato para todos los videos, se ubicará la cámara siempre en el mismo lugar y con los mismos estereotipos de vehículos, es decir, automóviles con volante a la izquierda, con parabrisas trasero, de circulación cotidiana en la ciudad.

No se tomaron en cuenta situaciones en donde se involucre un volante del lado derecho debido a la falta de autos con dicha característica en el país, sin embargo ese detalle no afectaría al análisis de los videos.

En relación a la posición del Sol en los videos, no se consideró la misma, ya que hablando con los conductores, se llegó a la conclusión de que no había muchas situaciones donde el Sol les influyera la vista en la ciudad. Y de influirles, los conductores afirmaban que su visión mejoraba mediante la utilización del dispositivo para-sol.

En cuanto a los conductores, se tratará de sujetos tanto masculinos como femeninos (4 masculinos, 3 femeninos), de 18 a 30 años, con estaturas desde 1,70 mts a 1,90 mts donde no se tomaron situaciones con gafas oscuras.

Como limitación, se puede decir también que se pretende la construcción de un sistema desarrollado con c++ para uso exclusivo de PC, con las correspondientes librerías a utilizar para el procesamiento digital de imágenes, y queda exento el escenario de desarrollo de un sistema embebido para integrar a la computadora de un vehículo el cual podría ser una extensión a futuro del presente proyecto.

A continuación se comenzará mostrando la descripción de la base de datos de video, ya que al tratarse de un sistema con entradas, transformaciones, y salidas, se vio adecuado para la estructura del informe.

- Capítulo II –

2. Diseño de la base de datos.

Dentro del ámbito del sistema de detección de pérdidas de la atención, lo que se pretende en primera instancia es evaluar a partir de cierta cantidad de elementos de análisis, las distracciones que sufren los conductores a la hora de la conducción en distintos momentos del día. Así, se diseñará y creará una base de datos conformada principalmente por videos, los cuales servirán como entradas al sistema con el fin de analizar las conductas de los conductores y a raíz de las mismas obtener conclusiones pertinentes.

Esta base de datos tendrá ciertas características puntuales detalladas a continuación, las cuales serán respetadas por cada uno de los videos con la intención de lograr una colección de elementos con características similares.

2.1 Posición de la cámara y especificaciones.

2.1.1 Posición

Se llevarán a cabo filmaciones siempre con una misma cámara, colocada siempre en el mismo lugar con el fin de mantener homogeneidad en la totalidad de las filmaciones. Para esto, se revisaron las bases de datos que había disponibles en la web, y en base a las mismas se escogió dicha posición. No fueron incluidas en los análisis ya que las mismas o eran muy pequeñas o no eran compatibles con las librerías a utilizar.

Como segundo punto a tener en cuenta, se planea la utilización de la librería Dlib para el análisis de los videos, la cual utiliza marcadores faciales con el fin de ir obteniendo los distintos gestos y rasgos faciales de los individuos. (Ver figura 2.1) [6].

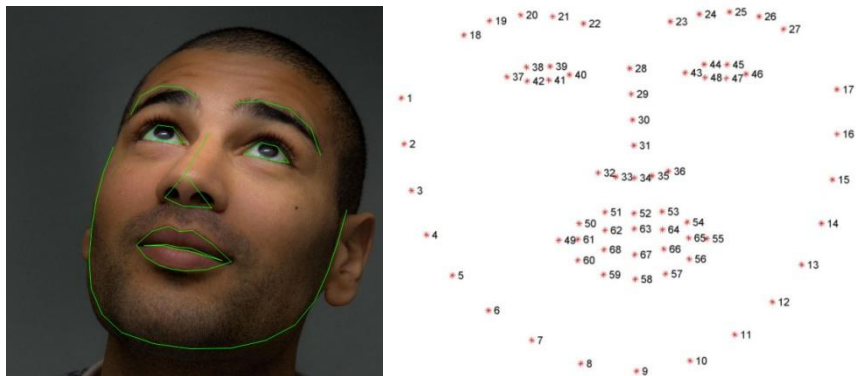


Fig. 2.1: Marcadores faciales brindados por Dlib

Mencionado lo anterior, se escogerá la posición de la cámara con la ayuda de las bases de datos que se mencionan a continuación, y con las características requeridas por Dlib, ya que la librería demanda sujetos de prueba en primer plano para la realización de una detección óptima [7].

En cuanto a las bases de datos analizadas, se puede decir que se relevó el estado del arte del proyecto a partir de 3 datasets provistos por la web, donde a partir de imágenes y videos se pudo verificar los distintos conjuntos de datos que había disponibles para este tipo de estudio.

Para dar a conocer un poco más dichos elementos, se hace mención al primer conjunto de estudio el cual consiste en 4 videos, donde frame a frame se van viendo las distintas actitudes de varios conductores (2 mujeres y 2 hombres) (Ver figura 2.2). Este dataset fue desarrollado por la empresa ElektraAutonomous Vehicle Developed la cual intenta lograr el desarrollo de autos con conducción autónoma, es decir sin la necesidad de interacción humana [8]. Este set de datos dio la iniciativa para crear una base de datos heterogénea con videos tanto de hombres como mujeres.



Fig.2.2: Conductores del primer dataset.

El segundo grupo de elementos está integrado por varios videos de hombres y mujeres y es un conjunto de datos robusto, provisto por el sitio Kaggle, el cual busca predecir 10 tipos de distracciones. (Ver figura 2.3). A diferencia del conjunto anterior, éste dataset tiene la cámara posicionada en un lugar poco óptimo, ubicada en la puerta del

acompañante del automóvil, haciendo que sea indistinguible la mitad de la cara del conductor [9].

Este set de datos es poco compatible con la librería Dlib, y se muestra solo para hacer mención a los tipos de datos que se encuentran disponibles en la web en el campo de las distracciones.



Fig. 2.3: Conductores del segundo dataset.

Para finalizar, el último dataset analizado en el estudio cuenta solamente con dos videos de muestra (Ver figura 2.4), los cuales sirvieron mayormente como guía para saber dónde ubicar la cámara dentro del automóvil. La pequeña base de datos fue provista por el grupo de investigación RobeSafe Research [10], y es menester mencionar que dicho grupo, utilizó estos videos para monitorear los comportamientos de los conductores utilizando la misma biblioteca que se pretende en éste proyecto (Dlib), lo que lo convierte en un dataset de gran interés.



Fig. 2.4: Conductores del tercer dataset.

Las bases de datos analizadas sirvieron exclusivamente de inspiración para la formación de la base de datos propia que utilizó el sistema y no fueron incluidas en los análisis de distracciones. Esto fue así porque los dos primeros conjuntos no respondían bien a la detección de rostros del sistema, y el tercer conjunto de datos se consideró muy pequeño para el análisis.

Se encontraron y se analizaron distintas bases de datos las cuales utilizaron variadas ubicaciones para su cámara (Ver figura 2.5). Sin embargo, teniendo en cuenta las

expresiones faciales requeridas, se propone posicionar la cámara en el tablero de velocidades.

Esta ubicación y la ubicación en el elemento “para-sol” del auto fueron las que mejor resultado trajeron en la detección de rostros. Finalmente se escogió la posición del velocímetro ya que en una posible implementación futura en un automóvil, mediante una cámara pequeña, evitaría que el conductor se distraiga y tampoco bloquearía su visión del tablero.



Fig. 2.5: Distintas posiciones de cámara encontradas en las bases de datos

2.1.2 Especificaciones

Como ya se ha mencionado, los videos serán capturados con un mismo dispositivo que en principio sería una cámara pequeña estilo GoPro con el fin de no bloquear la visión del conductor en ningún momento, sin embargo, se optó por otra alternativa más económica pero de igual calidad la cual refiere a una cámara brindada por un teléfono móvil sujeta a un soporte totalmente especializado para la superficie de interés (Ver figura 2.6).

El teléfono en sí tiene dimensiones grandes las cuales impiden ver de forma correcta el velocímetro del auto, sin embargo será utilizado en este proyecto con el fin de grabar los videos. En una futura implementación dentro de un auto la cámara sería colocada de una forma discreta en la posición mencionada.



Fig. 2.6: Posición de cámara desde diferentes ángulos.

En cuanto a las especificaciones técnicas del dispositivo, se destaca un sensor de 16 mega pixeles desarrollado por la empresa OmniVision (OV16860) el cual viene a ser uno de los sensores fotográficos de tope de gama en lo que respecta a smartphones [11]. Sus características más importantes radican en su capacidad de capturar vídeo 4K (3840 x 2160p) a un framerate de 60 imágenes por segundo, lo que le permite capturar imágenes en Ultra Alta Definición y realizar videos de 60 FPS.

2.2 Características generales de los videos

2.2.1 Datos técnicos

Se ha elegido un formato MP4 con resolución 1080p con 1920 x 1080 pixeles, cuya definición se vio indicada para el procesamiento de imágenes ya que la alta resolución de los videos permitió la obtención de frames con mayor detalle.

En cuanto a la duración y características del video, se simula la conducción de un vehículo durante un minuto aproximadamente. Donde el sujeto de prueba realizará una distracción cada 15 segundos (3 distracciones por video) indicada por un supervisor. El cual será el mismo que de inicio a la grabación del video, así como también será responsable de llevar el tiempo con un cronómetro, de modo que el sujeto de prueba se limite solamente a la simulación de la conducción (Ver figura 2.7).

Se eligieron simulaciones en lugar de videos reales de conducción, ya que el teléfono con el que se grabaron los videos ocupaba gran cantidad de lugar en el velocímetro y no permitía la correcta visión del mismo, lo que hace a la conducción una actividad

poco segura. Además se eligió videos de un minuto con la intención de hacer varios videos distintos y no ocasionar pérdidas de tiempo a los sujetos de prueba.

En los momentos de vigilia sin distracción, el sujeto será libre de realizar las acciones simples de una conducción correctamente realizada (mirar hacia las direcciones de las esquinas o simplemente tener la vista en el camino en todo momento)

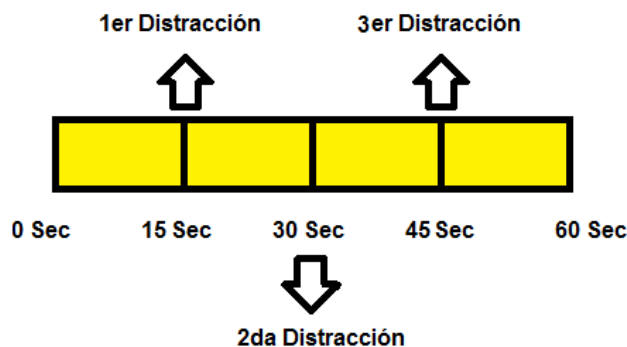


Fig. 2.7: Duración de los videos y distracciones en los mismos

2.2.2 Características generales

Los conductores que aparecen en los videos, o también llamados sujetos de prueba en este informe, son tanto masculinos como femeninos (4 masculinos, 3 femeninos), de 18 a 30 años, con estaturas desde 1,70 mts a 1,90 mts, y los mismos se ubican en forma estándar dentro del automóvil (No hay situaciones con los asientos reclinados, o situaciones que no se vean en la cotidianeidad)

Para lograr una base de datos más robusta, se realizaron videos a lo largo del día, realizados a la mañana, a la tarde, y en momentos nocturnos. Para cada situación mencionada se eligió el mediodía, las 5 PM, y las 9 PM con el objetivo de que se vean las diferentes incidencias de luz en los correspondientes videos.

Para continuar con las características generales de los videos, se definirán a continuación las diferentes distracciones con las que se trabaja en el sistema, llamando así Distracciones Leves, a aquellas que provocan la pérdida de la atención del conductor en el período aproximado de 1 a 3 segundos. Distracciones Moderadas, serán aquellas que provocan la pérdida de la atención en el conductor en el período de 4 hasta 6 segundos inclusive y finalmente las Distracciones Peligrosas refieren a aquellas que provocan pérdidas que van a partir de los 6 segundos. Esta definición de distracciones fue planteada exclusivamente para el proyecto y es necesario aclarar que en la vida real las clasificaciones pueden variar.

Se clasificarán los videos de acuerdo a las distracciones mencionadas y a la hora de evaluarlas se revisará también el tipo de comportamiento que realiza el conductor, siendo distracciones donde el sujeto pierde la visión de forma total en el camino, o si solo pierde la visión de forma parcial en el camino. Estas definiciones se profundizarán más adelante.

Una vez definidas las distracciones, se planteó un esquema tanto para una base de datos de videos de entrenamiento como una de prueba. Con la base de datos de entrenamiento se trabajó para ajustar los parámetros que componen al sistema y con la base de datos de prueba se obtuvieron los resultados que se presentarán en la sección 4. A continuación se deja la descripción de los esquemas junto con los datos de los sujetos:

Base de datos: videos de entrenamiento

- Distracciones leves
 - Video a la mañana (12 hs)
 - Video de hombre (24 años)
 - Video de mujer (21 años)
 - Video a la tarde (17 hs)
 - Video de hombre (20 años)
 - Video de mujer (30 años)
 - Video a la noche (21 hs)
 - Video de hombre (18 años)
 - Video de mujer (21 años)
- Distracciones moderadas
 - Video a la mañana (12 hs)
 - Video de hombre (24 años)
 - Video de mujer (21 años)
 - Video a la tarde (17 hs)
 - Video de hombre (20 años)
 - Video de mujer (25 años)
 - Video a la noche (21 hs)
 - Video de hombre (18 años)
 - Video de mujer (21 años)
- Distracciones peligrosas
 - Video a la mañana (12 hs)
 - Video de hombre (24 años)

- Video de mujer (21 años)
- Video a la tarde (17 hs)
 - Video de hombre (19 años)
 - Video de mujer (21 años)
- Video a la noche (21 hs)
 - Video de hombre (18 años)
 - Video de mujer (21 años)

Base de datos: videos de prueba

- Distracciones Leves
 - Video a la mañana (12 hs)
 - Video de hombre (20 años)
- Distracciones Moderadas
 - Video a la mañana (12 hs)
 - Video de hombre (20 años)
- Distracciones Peligrosas
 - Video a la mañana (12 hs)
 - Video de hombre (20 años)
 - Video de hombre (20 años)
 - Video a la noche (21 hs)
 - Video de hombre (21 años)

Dicho esquema proporciona una base de 18 videos de entrenamiento y 5 videos de prueba (3.5 GB total) con los cuales se trabaja a lo largo del proyecto. A su vez, cada video es nombrado de acuerdo al momento y el tipo de distracción que está exponiendo, es decir, si buscamos el archivo de video de una Distracción Leve, a las 5 PM, proporcionado por un hombre, en la base de datos se buscará DL_T_V.mp4 (haciendo referencia a Distracción Leve, Tarde, Varón).

Se eligió este esquema ya que abarcaban los distintos momentos del día, los distintos tipos de distracciones y los distintos tipos de conductores (Masculino/Femenino). Además como ya se mencionó se creó un conjunto de datos de prueba el cual será detallado en la sección 4.5.1.

Finalmente, cada uno de los elementos, tendrá vinculado un archivo de texto (Ver figura 2.8) el cual detallará las acciones que se realicen en el video. Se detallan con precisión de segundos el momento y la duración de cada una de las distracciones en cada video, y a su vez, se deja explícito el tipo de comportamiento que se está contemplando, pudiendo ser distracciones donde el conductor pierde toda la atención

en el camino, o pérdidas de atención donde tiene una visión parcial del mismo (Ver figura 2.9 y 2.10).

Los tipos de distracción se detallarán en la Sección 4, sin embargo para entender mejor la ficha de los videos, se deja una tabla resumen que explica la clasificación de las distracciones.

Clasificación por tiempo	Clasificación cualitativa
Distracciones Leves	Pérdida total del camino
Distracciones Moderadas	Pérdida parcial del camino
Distracciones Peligrosas	

Tabla 2.1: Clasificación de distracciones por tiempo y cualitativamente

Así, una distracción leve, moderada o peligrosa, puede deberse a una pérdida total o parcial del camino. Se harán análisis de ambos factores y se emitirán resultados en las secciones siguientes

Algo que es menester destacar, es que a los sujetos de prueba se les indicó el tipo de distracción que debían hacer, siendo estas leves, moderadas o peligrosas, sin embargo nunca se les planteó de qué modo debían hacerlas. De modo que podían perder la vista en el camino de forma total o parcial.

A su vez dentro de las distracciones peligrosas se consideró también escenarios de **somnolencia** la cual será mejor explicada en la sección 4.3. Para la misma se destinaron 3 de los 6 videos de distracciones peligrosas. Por tanto cuando se habla de distracciones peligrosas se refiere tanto a distracciones en donde el conductor pierde la visión en el camino de forma total o parcial por más de 6 segundos debido a los movimientos de su cabeza, o por estar sufriendo un escenario de somnolencia en el cual comienza a entrecerrar los ojos por un tiempo prolongado.

Además, se registran en el archivo de texto la duración y el momento de cada episodio con la intención de obtener medidas de error una vez implementado el sistema. Finalmente, se adjunta una tabla con las características resumidas de la base de datos (Ver tabla 2.2).

Video: DL_T_V

Tipo de distracción: LEVE.

Tamaño archivo: 129 MB.

Edad conductor: 20 años.

Duración total del video: 01 minutos 03 segundos.

Duraciones Distracciones:

Primera distracción: DESDE 00 minutos 18 segundos HASTA 00 minutos 20 segundos.

Segunda distracción: DESDE 00 minutos 33 segundos HASTA 00 minutos 36 segundos.

Tercera distracción: DESDE 00 minutos 48 segundos HASTA 00 minutos 51 segundos.

Detalles Distracciones:

Primera distracción: Pérdida de la atención con vista parcial del camino.

Segunda distracción: Pérdida total de la atención sin percepción del camino.

Tercera distracción: Pérdida total de la atención sin percepción del camino.

Fig. 2.8: Contenido del archivo de texto que acompaña a cada video



Fig. 2.9: Ejemplos de pérdida de atención total sin percepción del camino

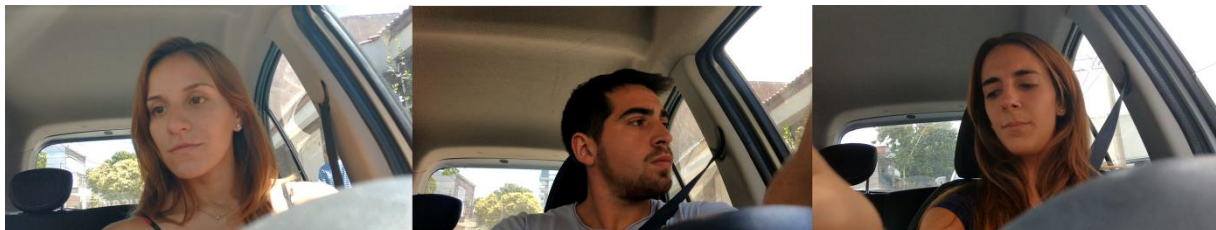


Fig. 2.10: Ejemplos de pérdidas de la atención con vista parcial del camino

Cantidad de videos	18
Cantidad de videos hombre	9
Cantidad de videos mujer	9
Cantidad participantes hombre	4
Cantidad participantes mujer	3
Tamaño base de datos (GB)	2,25
Tamaño promedio de archivos (MB)	130
Resolución de los videos (mega pixeles)	1920 x 1080
Velocidad del fotograma	28 fotogramas/segundo
Duración total de videos	18 Minutos 43 Segundos
Duración total de distracciones LEVES	0 Minutos 51 Segundos
Duración total de distracciones MODERADAS	1 Minutos 36 Segundos
Duración total de distracciones PELIGROSAS	2 Minutos 3 Segundos
Cantidad de distracciones	54
Cantidad de distracciones falta de visión PARCIAL	28
Cantidad de distracciones falta de visión TOTAL	26
Duración total distracciones v. PARCIAL	2 Minutos 7 Segundos
Duración total distracciones v. TOTAL	2 Minutos 19 Segundos
Duración total conducción sin distracción	14 Minutos 17 Segundos

Tabla 2.2: Tabla resumen de base de datos de videos

2.2.3 Metodologías para el registro de videos

Se detallan a continuación los pasos necesarios para la grabación de los videos:

- 1) Colocar la cámara en el soporte global siguiendo las instrucciones del fabricante del soporte
- 2) Situar en el asiento del conductor al sujeto de prueba
- 3) Realizar videos de prueba para corroborar la correcta posición del conductor (se deberá ver el rostro del sujeto en todo momento)
- 4) Una vez corroborada la posición, dar comienzo al primer video, controlando con un cronómetro la duración del mismo.
- 5) Cada 15 segundos, indicar al conductor que realice una de las distracciones de acuerdo al tipo de video que se esté realizando.
- 6) Finalizar el video cuando culmine el minuto.

2.3 Otras consideraciones sobre la base de datos

Para dar ejemplo de los videos capturados, se procede a mostrar una serie de capturas.

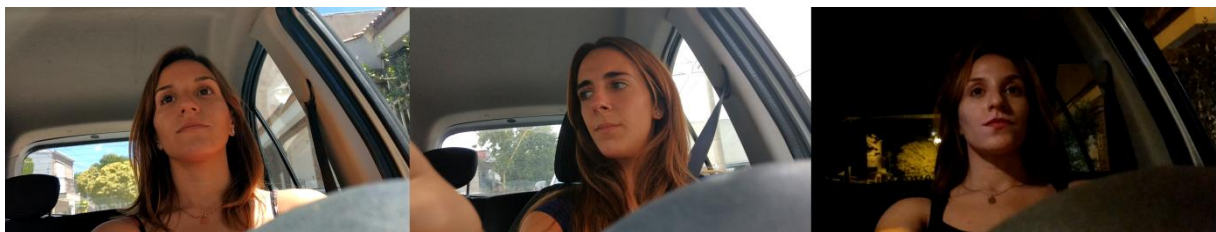


Fig. 2.11: Capturas de distracciones leves a lo largo del día para mujeres.



Fig. 2.12: Capturas de distracciones Moderadas a lo largo del día para varones.



Fig. 2.13: Capturas de distracciones peligrosas a lo largo del día para ambos sexos.



Fig. 2.14: Capturas de distracciones en videos de prueba.

En primera instancia se llevó a cabo la grabación de videos en donde quedaron en evidencia las diferentes distracciones que pueden ocurrir en la conducción. Se grabaron los mismos en diferentes momentos del día, y se llegó a la conclusión de que los videos tanto de la mañana como los de la tarde, presentaron características óptimas para ser tratados, mientras que los videos nocturnos tuvieron que ser grabados en calles con buena iluminación, debido a que de otra forma no se habrían podido grabar de forma exitosa. Cabe mencionar que éstos últimos, recibieron a su vez la ayuda de la luz auxiliar del automóvil, por lo que es un punto a resaltar a la hora de hacer los futuros análisis en dichos videos.

- Capítulo III –

3. Desarrollo del sistema

En esta etapa, se genera un sistema que recibe como entrada un flujo de video, el cual es procesado con el fin de detectar las pérdidas de atención. Luego se implementa un sistema basado en alarmas, en el cual se analizarán los videos a través de varios métodos y con un pre-procesamiento llamado Corrección Gamma.

Se implementó la detección de distracciones en los distintos videos, y se crearon puntos de control con el fin de alertar al usuario de los diferentes comportamientos incorrectos que está cometiendo el conductor.

3.1 Sistema operativo y librerías

Antes de comenzar con el desarrollo de los distintos métodos, se hace menester realizar una pequeña mención a las herramientas que se eligieron para la implementación del sistema. Así, se menciona en primer lugar Ubuntu 16.04, la biblioteca Opencv 3.4.2 la cual fue utilizada para todo lo que respecta al procesamiento de imágenes, y finalmente la biblioteca Dlib, que fue seleccionada para realizar la mayor parte del reconocimiento de las partes del rostro para luego utilizar dicha información en la detección de distracciones.

3.2 Diseño general del método

Se presenta el diagrama de bloques que se definió para el sistema y que permite observar la secuencia de pasos para realizar detecciones. Se pasa a dar detalle sobre los diferentes bloques que constituyen el sistema (Ver Figura 3.1).

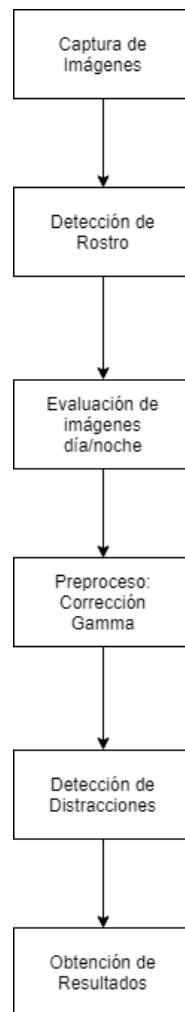


Fig. 3.1: Diagrama de bloques del sistema.

Como se puede observar en el diagrama, el primer elemento del mismo corresponde a la captura de imágenes la cual ya se vino explicando en secciones anteriores.

A continuación se explica las etapas del desarrollo correspondientes a la detección del rostro, clasificación de imágenes y pre-procesamiento, las cuales fueron desarrolladas con la ayuda de librerías open source como OpenCV y Dlib. Las últimas instancias del sistema serán desarrolladas en las secciones futuras.

3.2.1 Detección de Rostro

Para la detección del rostro del conductor, OpenCV utiliza dos metodologías llamadas clasificadores en cascada Haar y patrones binarios locales (LBP) [12]. Se probaron implementaciones con ambos algoritmos y se escogió el que mejor se acoplaba al proyecto.

Clasificadores en cascada Haar

En el caso de los clasificadores en cascada Haar, el método pretende obtener una sub-imagen a la que se le llama “imagen integral” a partir de la imagen de interés (en este caso la imagen del conductor). Para construir esta imagen integral, el algoritmo usa rectángulos blancos y negros a los que se les llama “Descriptoros” (Ver figura 3.2), los cuales se distribuyen a lo largo de la imagen principal obteniendo las intensidades de los distintos pixeles [13].

A esta imagen integral se le aplica un algoritmo llamado Adaboost [14], el cual divide la imagen integral en subregiones de tamaños diferentes y utiliza una serie de “clasificadores en cascada”, cada uno con un conjunto de características visuales con el fin de ir descartando aquellas secciones que no pertenezcan al rostro.

Para lograr esto los clasificadores se basan en premisas, tales como que la sección de los ojos es más oscura que la zona de la nariz y las mejillas, o que los ojos en sí son más oscuros que el puente de la nariz. Luego de éste último paso se obtiene la imagen final [15].

Patrones Binarios Locales

Lo destacable del método de patrones binarios, es que se trabaja pixel a pixel, evaluando la vecindad de los mismos (Ver figura 3.3)[16]. Se analizan los 8 vecinos de cada pixel, y se van evaluando las intensidades de los vecinos. Si un vecino es más intenso que el pixel en cuestión, se asigna un 1, y si es menos intenso, un 0.

Como se ve en la Figura 3.3 luego de analizar los vecinos, se convierte el número binario obtenido en un número decimal que corresponderá a un nuevo valor de intensidad de la imagen. Luego se divide la imagen en regiones, y se calcula un histograma para cada una de ellas. Es así que a través del análisis de estos histogramas el método contribuye a la detección de rostros [17].

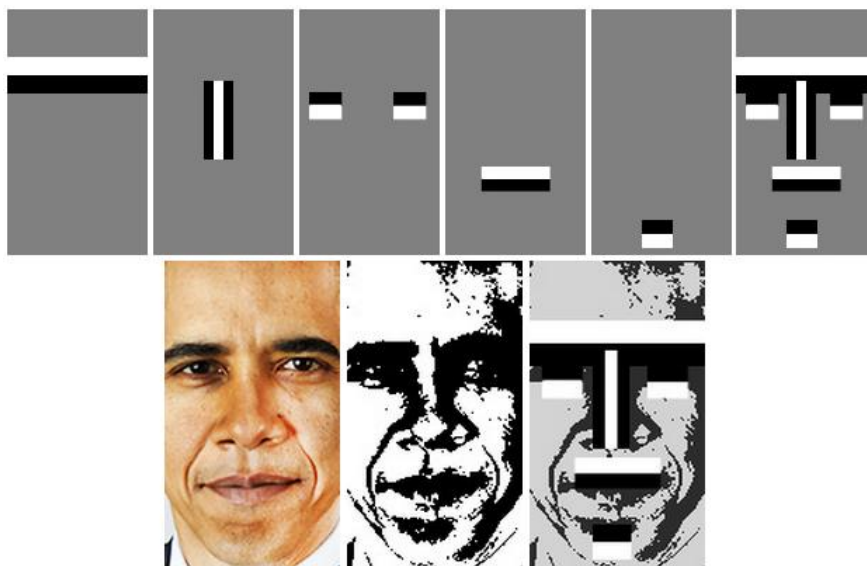


Fig. 3.2: Método de detección de rostros por clasificadores en cascada

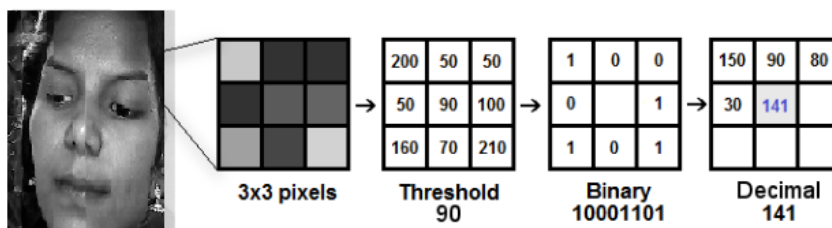


Fig. 3.3: Ejemplo de vecindad en el método de Patrones Binarios Locales

Con el método de OpenCV de detección facial, es posible obtener la localización del rostro del conductor, para luego con Dlib [18] obtener los 68 puntos característicos del mismo (Ver figura 3.4), tal como muestra la Figura 3.4. Estos 68 puntos serán los que den la información de la posición de la cabeza del conductor en todo momento, y se explicará con mayor detenimiento en la sección 3.2.2.

Para la detección de rostros propiamente dicha, se optó por utilizar el método de clasificadores en cascada Haar provisto por OpenCV, ya que mediante varias pruebas, se llegó a la conclusión que era más óptimo que el método LBP.

Para llegar a esto, se realizaron 18 pruebas, una por cada video donde se evaluó la cantidad de frames sin detectar el rostro por ambos algoritmos, ya que fue de suma importancia que el software detecte correctamente cuando el conductor estaba en momentos de vigilia. Se analizaron los videos y corriendo los algoritmos se verificó cual

de los algoritmos detectaba correctamente el rostro cuando el conductor conducía atento al camino.

Para ver si el sistema detectaba o no una cara, OpenCV crea un vector en cada frame llamado *Faces* para verificar si detecto o no el rostro del conductor [19], lo que se hizo fue chequear que ese vector tenga incluida la cara del conductor. Si no detectó la misma, el vector estará vacío.

Del número total de pruebas, el detector basado en LBP logró una tasa de acierto del 27% en relación al detector basado en Haar, el cual tuvo resultados del 73% tal como se aprecia en la Tabla 3.1. Cabe mencionar que ambos métodos tienden a números más elevados en los videos nocturnos debido a la falta de luz. Dichos porcentajes muestran frames acumulados a lo largo del vídeo, en los cuales el conductor tendría que estar conduciendo correctamente.

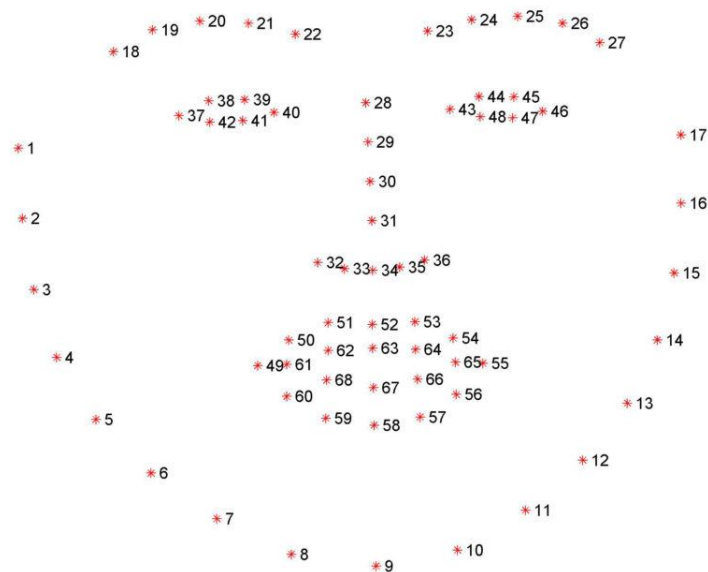


Fig. 3.4: Marcadores faciales obtenidos mediante la biblioteca Dlib.

Video	Frames sin detectar (Haar)	Frames sin detectar (LBP)
DL_M_M	12	18
DL_M_V	7	14
DL_T_M	11	9
DL_T_V	20	27
DL_N_M	100	100
DL_N_V	95	99
DM_M_M	70	89
DM_M_V	75	70
DM_T_M	67	80
DM_T_V	55	70
DM_N_M	100	99
DM_N_V	93	96
DP_M_M	114	110
DP_M_V	105	124
DP_T_M	99	115
DP_T_V	98	125
DP_N_M	130	160
DP_N_V	129	150

Tabla 3.1: Resultados de comparación entre detector basado en Haar y detector LBP.

En la tabla se aprecian diferencias muy pequeñas entre los clasificadores en cascada y LBP, como mucho de 30 frames acumulados en todo el video, sin embargo esto fue realizado para la optimización del sistema y serviría en caso de videos más extensos.

3.2.2 Estimación de postura inicial

Aquí se explica detalladamente cómo se realiza la estimación de la profundidad en el rostro del conductor y la mirada del mismo en la escena analizada. Para detectar las actitudes de pérdida de la atención, el sistema debió ser capaz de percibir los diferentes movimientos del sujeto, basándose en las diferentes posiciones de la cabeza. Una vez detectados los movimientos se procedió a una evaluación de los mismos con el fin de tomar decisiones.

Como se mencionó en párrafos anteriores, la biblioteca Dlib aportó la información necesaria sobre las posiciones de 68 puntos de interés en el rostro humano, sin embargo, lo que es de importancia señalar, es que dichos puntos se encontraron en un plano de dos dimensiones por lo que cada parte del rostro viene con ubicación X-Y, despreciando la profundidad. Así, surgió el principal problema de esta primera parte ya que es necesario obtener la coordenada de profundidad para evaluar la posición de la cabeza dentro de los videos de entrada. Para dar solución a éste problema, se realizó

un mapeo del plano de dos dimensiones, a un espacio de tres dimensiones (Ver figura 3.5), siempre basándose en las posiciones de Dlib. Para una aceleración de cálculos, se utilizaron solo algunos de los 68 puntos (14 en cuestión detallados a continuación), los cuales eran suficientes para hacer un mapeo realista.

En cuanto a los puntos del modelo tridimensional, ya que es necesario que cada punto de la biblioteca se mapee con un punto del espacio 3D, fueron obtenidos de la web [20] junto con la imagen presentada a continuación y cada uno de los puntos correspondientes a partes del rostro tales como la esquina del mentón, las esquinas de la nariz, las esquinas de los ojos, y puntos relacionados con las cejas (es decir, los 14 puntos de los que se habló previamente)

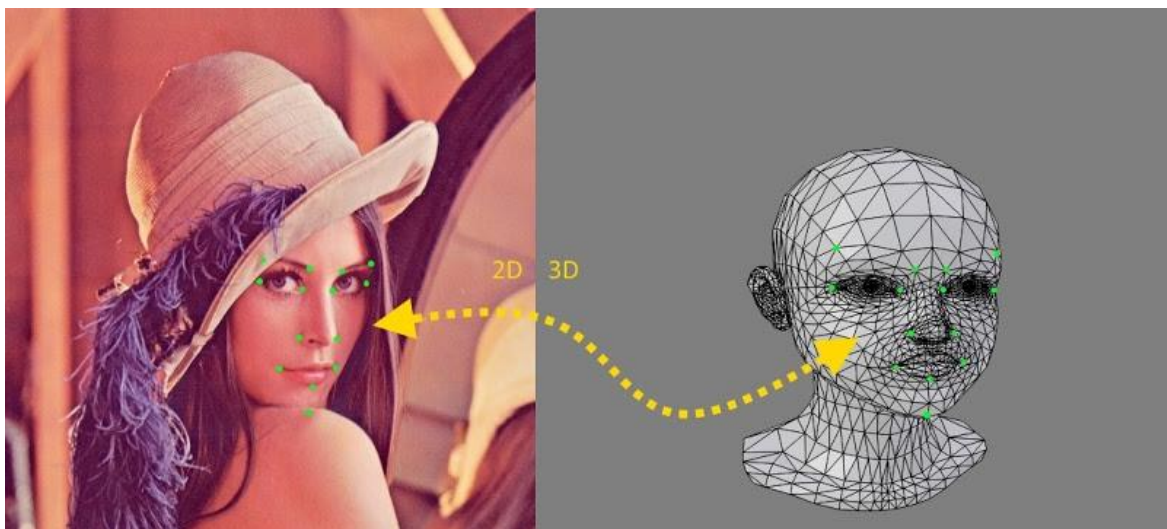


Fig. 3.5: Mapeo del conjunto de 14 puntos 2D a un modelo 3D

Una vez definidos los puntos del modelo 3d, mediante funciones de Open CV se calcula el centro de dichos puntos, y en dicho centro se traza un eje de coordenadas fijo, el cual servirá para calcular los ángulos en los que se mueve la cabeza del conductor.

Imagen tras imagen se van calculando los distintos ángulos de rotación de la cabeza del sujeto en relación al eje fijo que se trazó previamente. Los ángulos de un sistema de rotación en relación a un sistema de coordenadas fijo, son llamados ángulos de Éuler [21] y son brindados por la librería Open CV en base a los puntos del modelo 3d que se van moviendo a lo largo del video.

Algo que vale la pena aclarar de ahora en más, es que, dentro de los ángulos de Euler, se utilizarán los ángulos de navegación: Yaw – Roll – Pitch, los cuales son los que mejor describen el movimiento de un objeto de 3d móvil respecto de un eje.

En otras palabras, los valores del ángulo Roll muestran la rotación del eje móvil X con respecto al eje fijo X, los valores del ángulo Pitch muestran la rotación del eje móvil Y con respecto al eje fijo Y, y los valores del ángulo Yaw reflejan las rotaciones del eje Z [22]

En la figura se puede apreciar como irían variando los ángulos $-\theta(Yaw)$, $\psi(Roll)$, $\varphi(Pitch)$ -, teniendo un sistema de ejes fijo -azul-, y un sistema de ejes que varía con el tiempo - Rojo - Asi como también se muestra el vector N, el cual se forma mediante la intersección de los planos formados(Ver figura 3.6)

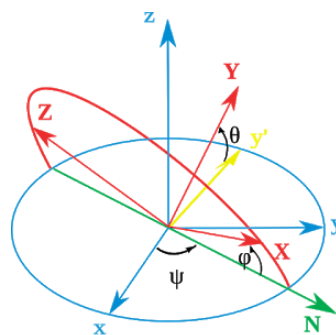
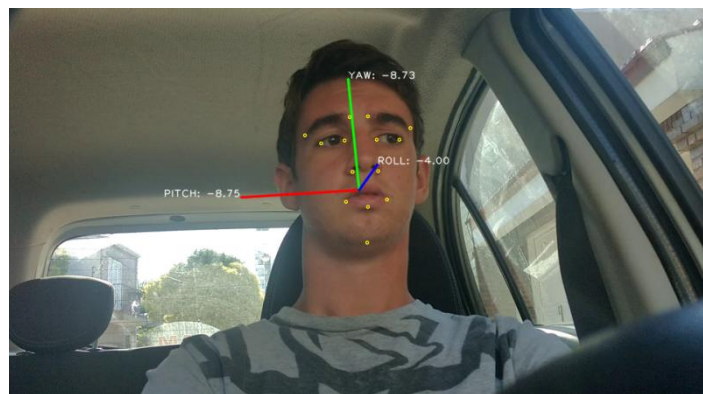


Fig. 3.6: Ejemplo de Detección utilizando ángulos de Euler y versión conceptual de los mismos.

3.2.3 Evaluación de imágenes día/noche

Como se planteó con anterioridad, el sistema toma como entradas una serie de videos con el fin de analizarlos y proceder a una conclusión pertinente. Hace falta mencionar que dentro de la base de datos, un tercio corresponde a videos de categoría nocturna, por lo que resultó necesario el desarrollo de un pre-procesamiento para tener un resultado más óptimo a la hora de analizarlos.

Para el pre-proceso se realizaron pruebas sobre los histogramas de cada entrada (Ver figura 3.7), donde se evaluó el primer segundo (es decir los 28 primeros frames) de cada video con el fin de dar una conclusión sobre si era nocturna, o diurna. Se utiliza dicha cantidad de frames ya que es lo que requiere la cámara para estabilizarse.

Cada frame fue evaluado y mediante una función llamada *DayOrNigth* se obtuvo un resultado Verdadero, si la imagen correspondía a un escenario de noche, y el resultado sería Falso, si la imagen correspondía a un escenario de día. Esta función utiliza los histogramas del frame analizado, es decir los gráficos asociados al frame que describen la distribución del color en la escena.

Los parámetros más importantes de esta decisión se dan en la evaluación del histograma. Donde la función retornará Verdadero, si el histograma AZUL de la región de interés en promedio, da menor a 90. Resultaba adecuado para caracterizar los videos con presencia de noche.

Para decidir si el video es de día o de noche, se procedió a evaluar dentro del bucle normal del programa, si la función mantiene su resultado para al menos 20 frames de los evaluados. El video queda catalogado como video nocturno y pasa a la siguiente instancia de su pre-proceso, caso contrario que no llegue a los 20 frames de acierto, será catalogado como video de día.

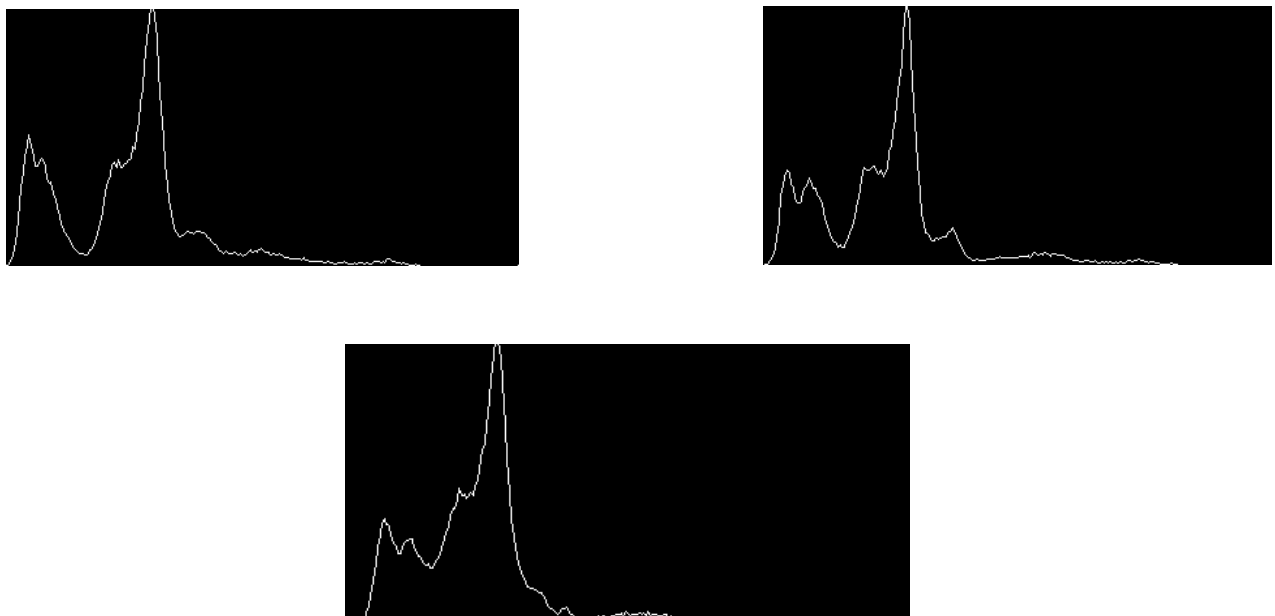


Fig. 3.7: Ejemplos de Histogramas RGB para una imagen de día. Imagen superior Izquierda: Histograma Azul (B). Imagen superior Derecha: Histograma Verde (G). Imagen Inferior: Histograma Rojo (R)

Mediante reiteradas pruebas, se llegó a la conclusión de que la evaluación debía realizarse sobre los histogramas del conjunto de colores RGB, ya que en ambientes con poca iluminación al ser colores aditivos se presentan valores mucho más pequeños tanto de rojo, verde o azul. Al ser los histogramas fuente de información sobre la distribución del color en una imagen, se investigó sobre qué parte evaluar para obtener las características del video. Para dar solución a esto, y pensando en que todos los videos serían filmados desde la misma posición, se planteó un ROI (Región of Interest – Región de Interés) (Ver figura 3.8) el cual representa una porción del frame evaluado, lo suficientemente significativo como para obtener una conclusión acertada sobre la condición del video.

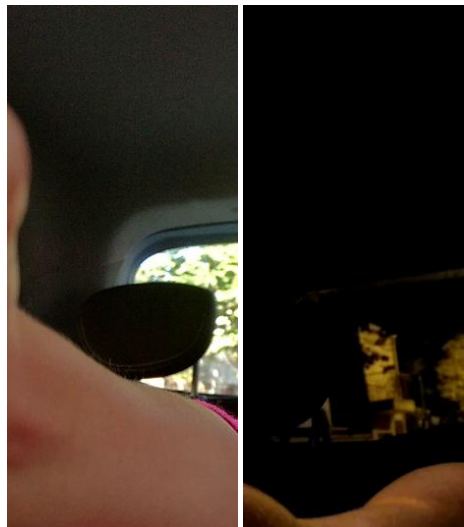


Fig. 3.8: Ejemplos de ROI tanto para una imagen de día y para una nocturna.

3.2.4 Corrección Gamma

Para el pre-proceso de los videos nocturnos, se pensó una transformación que pudiera resaltar el contraste de los videos, con la posibilidad de que mejore la detección de rostros de la biblioteca elegida, para esto se eligió la denominada transformación o corrección Gamma.

La corrección Gamma es una operación no lineal la cual se usa para codificar y decodificar luminancia o valores triestímulos en sistemas de video o imagen [23]. Así, definida por la ley de las potencias está descrita de la siguiente forma:

$$V_{out} = AV_{in}^{\gamma} \tag{1}$$

Para la simplicidad del método, se utilizó la forma común de la fórmula con un A constante igual a 1 y en cuanto al gamma, se fue aplicando con distintos valores hasta encontrar un resultado óptimo. (Ver Figura 3.9) Básicamente con éste método se buscó darle un mayor contraste a la imagen con el fin de mejorar la detección en el sistema.

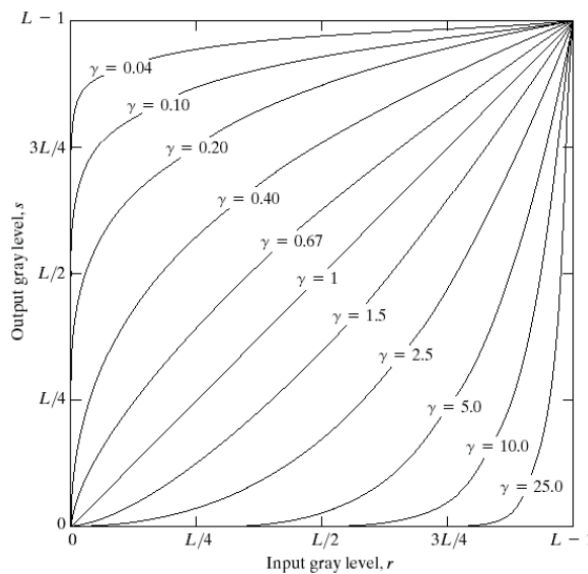


Fig. 3.9: Diferentes valores de Gamma junto con su curva logarítmica

Para evaluar el resultado óptimo, se tomó video por video viendo con qué factor gamma la biblioteca detectaba mejor los rostros. En esencia, la detección buena, o mala, estaba representada por la simple detección de la cara del conductor.

Es así que para evaluar el adecuado preprocesamiento de los videos se veía que tantos frames detectaba el sistema en situaciones de conducción correcta, donde cuanto más frames detectaba, mejor era el gamma.

Evaluando con valores desde 0.1 hasta 0.9, con un paso de 0.1 se llegó a la conclusión de que el mejor valor para el gamma correspondía al valor 0.6, por lo que cada video que sea catalogado como "Video Nocturno" debió pasar por su correspondiente corrección gamma con valor de 0.6.

Se llegó a este número mediante la evaluación de todos los videos de noche, y las detecciones utilizando el índice 0.6 fueron las más satisfactorias, es decir, dentro de todos los índices probados éste fue el que la detección de la mayor cantidad de frames a la hora de detectar al conductor conduciendo.

No se utilizó un gamma adaptativo debido a que el índice obtenido funcionaba de forma satisfactoria en todos los videos y su implementación significaba una desviación en los tiempos del proyecto. Para una implementación futura sería una de las cosas a implementar.



Fig. 3.10: Capturas nocturnas sin corrección Gamma y con corrección Gamma correspondientemente.

Hasta aquí se ha explicado el desarrollo que va desde la adquisición del video hasta encontrar la orientación de la cara del sujeto y los puntos que caracterizan su rostro. A continuación, se presentan las definiciones de los diferentes tipos de distracciones, como se realiza su detección y el sistema de alarmas.

- Capítulo IV –

4. Evaluación de distracciones

En este trabajo se definieron determinados aspectos físicos de sobre las conductas de las personas involucradas en la conducción de un automóvil. Con base en la observación de éstos, se definieron distintos factores (consideraciones más bien abstractas) que debían ser evaluados en los conductores para determinar si respectaban una conducta esperada durante la conducción, a saber, que no perdieran la atención del camino y que no presenten un grado de somnolencia inadmisibles para el manejo. Una vez escogida la localización de la cámara y con ello la escena, se establecieron los sistemas cartesianos de referencia y sobre éstos se establecieron los ángulos de la cabeza en los que las personas respetarían las condiciones de manejo y la velocidad de parpadeo admisible de forma ad hoc. Una vez establecidos los parámetros, se probó el sistema y obtuvieron resultados comparando con un conjunto de prueba.

Como se mencionó en el capítulo 2, además de las distracciones Leves/Moderadas/Peligrosas basadas en el tiempo en el que el sujeto se distrae, se creó otra clasificación basándose en la pérdida de visión del sujeto con respecto al camino. Esto se realiza para ver un análisis cualitativo a la hora de ver los resultados de las distracciones leves, moderadas y peligrosas.

De esta manera, se definen las Distracciones con pérdida total de la visión a aquellas distracciones que impliquen un comportamiento totalmente descuidado que lleve al conductor a abandonar la vista completamente del camino. Por ejemplo mirar hacia atrás o hacia los costados por un determinado período de tiempo.

Por otro lado, llamaremos Distracciones con pérdida parcial de la visión a comportamientos que si bien podrían calificarse como distracciones, no hacen perder la vista en su totalidad de manera que el conductor percibe aun una visión periférica del camino (Ver Figura 4.1).



Fig. 4.1: Diferencias entre distracciones Parciales y Totales.

4.1 Distracciones con pérdida de visión total

Para el caso de las Distracciones con pérdida total de la atención, se planteó una característica común a todos los videos, y era que el conductor al realizar éstos comportamientos, la biblioteca inmediatamente dejaba de detectar el rostro del sujeto. Esto puede sonar evidente, ya que al realizar una Distracción como la mencionada, el sujeto de prueba no tiene contacto alguno con el camino, y por lo tanto, con la cámara.

Es acertado pensar que un solo frame sin detectar un rostro no significa que el conductor se esté distraiendo, puede tratarse tranquilamente de un problema de la cámara por lo que para detectar efectivamente si el conductor está perdiendo la atención, se controló la ausencia de rostros en 28 frames seguidos. Es ahí donde se detecta 1 segundo de distracción causado por una distracción con pérdida de visión total. Luego se plantearán parámetros para el ajuste personalizado de estos tiempos, por ejemplo si se querían obtener distracciones de medio segundo, se bajaría ese rango de frames a la mitad.

El video de entrada se fue ejecutando y en tiempo real el sistema fue capaz de reconocer distracción, la cual de acuerdo al tiempo que dure, será catalogada como Leve – Moderada – Peligrosa. Para tener las ideas más claras, los tiempos asignados a cada una de las distracciones L.M.P son los siguientes:

- **Distracciones Leves:** Equivalentes a 1 a 3 segundos del conductor distraiéndose.
- **Distracciones Moderadas:** Equivalentes a 4 o 5 segundos del conductor distraiéndose.
- **Distracciones Peligrosas:** Equivalentes a 6 o más segundos del conductor distraiéndose. También está considerada distracción peligrosa el factor somnolencia que será explicado brevemente a continuación y desarrollado en la sección 4.3

Todos estos tiempos fueron consensuados por los sujetos de prueba mediante entrevistas en las cuales se les preguntó a cada uno que tiempo consideraban a una distracción leve, moderada o peligrosa.

Como se mencionó previamente, las distracciones peligrosas a su vez están compuestas por el factor de la somnolencia. A la hora de ejecutar el sistema, se irán detectando distracciones causadas por el movimiento de la cabeza del conductor en tiempo real, y cada 15 segundos se irá haciendo una revisión tanto de las distracciones por movimiento, como la somnolencia del mismo.

4.1.1 Buffer de distracciones

Existen casos donde el conductor se está distraendo, y en un instante el sistema podría volver a detectar el rostro del conductor por simple equivocación de Dlib. Así, si el sistema reconocería un rostro en el frame 26 de una distracción (donde las mismas vienen de a 28 frames), el sistema no registraría una distracción, cuando el conductor efectivamente lo está haciendo. El sistema en este caso estaría funcionando de forma incorrecta.

Para evitar casos donde el sistema detecta rostros de forma incorrecta se creó lo que en el proyecto fue denominado como buffer de distracciones. Éste buffer realiza una acumulación cada vez que el sistema deja de detectar un rostro (cuando hay distracción). Así, si el sistema detecta un rostro, el buffer lanzará la detección de una distracción, siempre y cuando el buffer haya estado acumulado de 24 a 28 frames. Es necesario recordar que 28 frames corresponden a un segundo de video, y el frame 24 fue obtenido en base a pruebas preliminares. Esto da la posibilidad de que las distracciones no vengan de a segundos enteros, sino de a segundos variables.

4.2 Distracciones con pérdida de visión parcial

Las distracciones con pérdida de visión parcial son las referidas a las pérdidas de atención que implican cierto porcentaje de la vista en el camino, se habla de acciones como encender la radio, la utilización de cigarrillos al conducir, teléfonos celulares, buscar objetos, entre otras.

Para las detecciones con pérdida de visión parcial se utilizaron las posiciones de la cabeza del conductor en todo momento, para esto fueron de ayuda los ángulos de Euler ya que de acuerdo a los valores que se iban obteniendo de cada frame se iban sacando conclusiones sobre si el conductor se estaba distraendo o no. Para detectar estas distracciones fue de interés la posición brindada por el eje Yaw, el cual refiere al

que se encuentra en la zona vertical de la cabeza. Se escogió dicho ángulo para el análisis debido a los cambios bruscos que obtenía el mismo en todo momento.

Para detectar la distracción se evaluó la posición del ángulo Yaw, y entrará en zona de distracción si éste se encuentra entre valores MENORES a -15° o MAYORES a 20° . Estos valores fueron los óptimos para detectar las distracciones ya que dentro de un ámbito reducido como es un automóvil, es difícil tener grandes campos de visión, por tanto dichos campos de visión están reducidos a un espectro dentro esos valores.

Como lo anterior, estos parámetros serán ajustables en el caso de que alguien necesite evaluar conducción en alguien que no deba distraerse en absoluto, o todo lo contrario, que pueda distraerse por demás. Como las distracciones anteriores, también serán impulsadas por 28 frames consecutivos de cumplimiento de la condición, sin embargo aquí no se planteó buffer de distracción ya que se consideró que una distracción parcial era detectada por la cámara en todo momento.

Para aclarar las distintas formas entonces de generar una distracción, se describen los 3 escenarios que pueden pasar para generar una detección de distracción:

- 1) Que no se detecten 28 frames consecutivos de un rostro, escenario donde se detecta una distracción con pérdida total de la visión.
- 2) Que se detecten 28 frames consecutivos de un rostro cumpliendo la condición del ángulo de Euler mencionado, donde se detecta una distracción con pérdida parcial de la visión.
- 3) Que en el frame 26 por ejemplo de una distracción por pérdida total de visión el sistema detecte un rostro de forma equivocada, en cuyo caso saldría el buffer de distracciones diciendo que hay una distracción con pérdida total de la visión.

4.3 Parpadeos y somnolencia

Más allá de las distracciones previamente analizadas, existe otro fenómeno que afecta a gran cantidad de usuarios, y muchas veces de forma fatal: la somnolencia. Para el análisis de este factor, se planeó en primer lugar definir una herramienta para la detección y contabilización de los parpadeos de la persona, y en base a la cantidad de parpadeos detectados en cierto tiempo obtener respuestas. Este fenómeno será analizado en paralelo y en caso de existir somnolencia, será considerada una distracción peligrosa.

Para la detección de parpadeos fue utilizado un concepto obtenido de la web denominado EAR – *Eye Aspect Ratio* [24] que está basado en una serie de detecciones utilizando la ya mencionada biblioteca Dlib. Éste concepto utiliza los pares

de coordenadas x-y que Dlib brinda en relación a los ojos del conductor (Ver Figura 4.2).

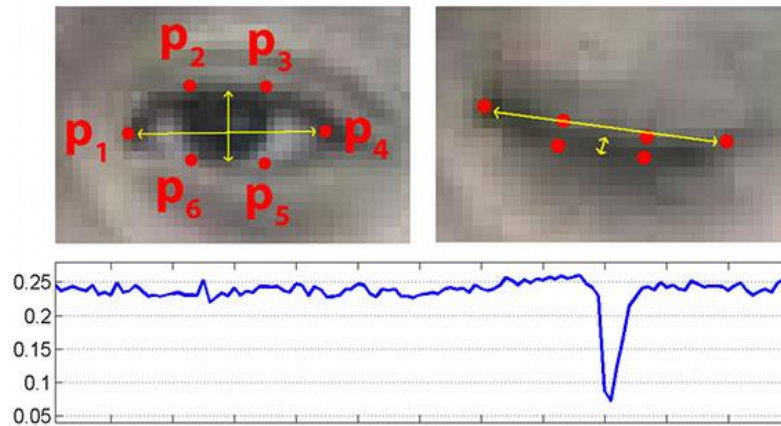


Fig. 4.2: Enumeración de puntos en Dlib

Como se ve en la siguiente fórmula, se intentó dar una relación entre las distancias euclídeas entre los puntos de interés. Se puede decir que el numerador computa la distancia entre los puntos verticales definidos por los marcadores, mientras que el denominador computa la distancia entre los puntos horizontales.

$$EAR = \frac{\|P2-P6\| + \|P3-P5\|}{2\|P1-P4\|} \quad (2)$$

Un dato a tener en cuenta, es que el número 2 que figura multiplicando en el denominador, le provee cierta ponderación al mismo, ya que la cantidad de puntos horizontales es menor que la de los puntos verticales. Lo interesante de este cálculo, más allá de todo procesamiento digital de imágenes, es que el EAR se mantendrá constante siempre que el ojo se mantenga abierto, y tenderá a cero al efectuarse un parpadeo.

Para la implementación del mismo, se realizó una estimación de 2 parámetros fundamentales: La cantidad de frames que corresponden a un evento de parpadeo, y el espectro deseado de EAR. Comenzando con el último ítem, se busca que el EAR tienda a cero, por lo tanto se define un umbral donde si el cálculo lo alcanza, será considerado parpadeo. A su vez, se nombro la cantidad de frames necesarios para que esto ocurra, entonces, se definió que si en determinados frames consecutivos, el EAR se encontraba por debajo del umbral deseado, se podía considerar parpadeo.

Aclarando un poco más los términos y yendo a los datos propios del sistema, se planeó un umbral de 0.25 para el EAR, que deberá ser cumplido en 3 frames consecutivos para ser considerado parpadeo. La cantidad de frames es pequeña ya que se trata de un evento fugaz.

Algo que es importante aclarar, es que en los videos de somnolencia el número de parpadeos se disparaba sin siquiera haber alguno, esto es porque los conductores comenzaban a entrecerrar los ojos de manera exagerada por lo que hicieron que el EAR comenzara a dar un índice diferente a los números de los parpadeos. Esto fue aprovechado, y el número de parpadeos se pasó a llamar **Índice de somnolencia**. Este índice varía según lo entrecerrado que tengan los ojos los conductores y en base a eso crece más o menos.

Como no se consideró situaciones que puedan encandilar al conductor, el entrecerrado de los ojos fue tomado como una acción de somnolencia, y no un reflejo resultante de tener el sol o una luz adelante.

Es así entonces que se evaluarán tanto las distracciones de los conductores por el movimiento de su cabeza, como el valor del índice de somnolencia, ya que si se presta atención, una persona podría estar con la vista fija en el camino, pero con los ojos totalmente cerrados.

4.3.1 Índice de somnolencia

Ante la posibilidad de un escenario de somnolencia, a medida que va avanzando el video se irá calculando el índice. El mismo se emitirá en base a los resultados de la formula descrita anteriormente junto con un mensaje relacionado descrito en la Tabla 4.1. Éste valor se calculó a lo largo de todo el video y se irá mostrando cada 15 segundos.

Índice de Somnolencia	Mensaje
1 – 25	Somnolencia controlada
26 – 50	Parpadeos Moderados
51 – 75	Atención Parpadeos excesivos
76 – 100	Peligro! Somnolencia Detectada
>100	Peligro! Alto Riesgo!

Tabla 4.1: Tabla de Índices de Somnolencia y sus mensajes

Todos los números presentados fueron establecidos en base al análisis de todos los videos y obtenidos de forma empírica. Se puede decir que los dos primeros ítems

coincidían generalmente a distracciones leves, los dos siguientes a distracciones moderadas y los últimos dos grupos están relacionados con distracciones peligrosas.

Para traducir los valores del índice a valores conocidos, se hizo una conversión del valor del índice a los segundos en los que el conductor estuvo bajo el efecto de somnolencia. Para esto se basó en la duración que tiene un parpadeo la cual es aproximadamente de 400 milisegundos [25], y ya que nuestro índice en primera instancia fue creado para la detección de parpadeos, se asume que 400 milisegundos equivalen a un 1 en el índice de somnolencia.

Índice de Somnolencia	Promedio en Segundos	Porcentaje del video	Mensaje
1 – 25	5,2	8,33%	Somnolencia controlada
26 – 50	15,2	25,33%	Parpadeos Moderados
51 – 75	25,2	42%	Atención Parpadeos excesivos
76 – 100	35,2	58,66%	Peligro! Somnolencia Detectada
>100	45,2	75,33%	Peligro! Alto Riesgo!

Tabla 4.2: Tabla de Índices de Somnolencia y sus respectivas equivalencias a segundos.

Para obtener los valores en segundos, se tomó un promedio de los índices de la fila y se multiplicó ese valor por 400 milésimas de segundo, luego se hizo la conversión de milisegundos a segundos. La columna “Porcentaje del video” muestra el porcentaje del video al que esos segundos pertenecen.

4.4 Diseño de alarmas temporales

4.4.1 Modelado de las distracciones

A partir de las metodologías presentadas hasta el momento, se planteó desarrollar una serie de puntos de control que sirvan de evaluadoras para los videos, de modo que al final de cada uno se pueda observar una conclusión del mismo.

Estos puntos de control son llamados “alarmas” ya que al momento de ver el análisis del video, junto al mismo irán apareciendo señalizaciones visuales que muestran el comportamiento que se analizó. Esto se profundizará en la sección 5.

Volviendo al formato original de los videos, es necesario recordar que éstos presentaban una estructura homogénea. Así, podemos mencionar que en estos videos de un minuto, había presentes momentos de distracción cada quince segundos, obteniendo así tres escenarios a detectar por el sistema.

Para comenzar con las alarmas es necesario previamente definir las estructuras de datos que se manejaron, debido a que dichas alarmas se activaron con algunas propiedades de dichas estructuras. Las distracciones al ser detectadas mediante los métodos previstos anteriormente, fueron modeladas en clases. Se creó una clase Distracción, la cual tendrá solo atributos públicos para una mayor facilidad de acceso a los mismos, donde estuvieron presentes el nro. de frame donde se detectó la distracción y el ID del tipo de distracción que se presentó. La segunda clase modelada, la cual se llamó TipoDistracción, que a su vez tenía también atributos públicos, donde fueron parte elementos como el ID del tipo de distracción, y una cadena de caracteres que explicaba el comportamiento (Distracción total o parcial). (Ver Figura 4.3).

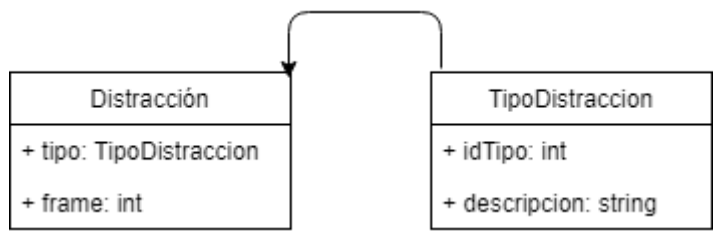


Fig. 4.3: Imagen simplificada de la relación entre clases

Dentro del sistema, se creó un vector del tipo Distracción, este vector fue el encargado de almacenar las distintas distracciones a lo largo del proceso. Una vez declarada la estructura, al detectar una de ellas, el sistema debió generar un elemento del tipo Distracción y colocarlo en el vector. A su vez es necesario recordar que con la creación de un objeto Distracción, se creaba a la par uno del tipo TipoDistracción.

Para generar una pérdida de la atención correspondiente a su clase, el sistema esta atento al escenario que ha detectado, y en base al mismo, genera un objeto del tipo distracción. Así, si estamos en presencia de una actitud de pérdida de visión Total, el sistema se encargaba de generar un objeto con el id correspondiente a dicha categoría. Lo mismo sucedió con pérdidas de visión parciales.

Al crear las instancias, el sistema también utilizó propiedades de OpenCV para adquirir el frame al cual hace referencia esa escena, la cual nos sirvió para hacer conclusiones pertinentes y brindarnos información sobre la dinámica temporal. Ya con todo lo mencionado, se puede hablar de un entorno el cual detectaba escenas deseadas, y las

ubicaba en una estructura de datos con su correspondiente tipo, y su determinada cronología.

4.4.2 Desarrollo de alarmas progresivas

Se pretende evaluar si el sistema fue capaz o no de poder reconocer los escenarios que se le presentaron, para lo cual se realizaron 3 distracciones presentadas en cada video, una cada 15 segundos.

Para lograr una correcta evaluación de la conducción, al igual que las distracciones, se crearon puntos de control cada 15 segundos, con el objetivo de ir evaluando los escenarios transcurridos hasta el momento. Se plantearon 4 puntos de control en total en los segundos 15, 30, 45, y un último punto antes de finalizar el video en el segundo 59.

El primer punto de control evalúa los primeros segundos del video, el segundo evalúa la segunda parte, el tercero la tercera parte, y el último punto de control evalúa el último tramo del video. Todas las alarmas son de características idénticas y lo que varió fue su disparo temporal. Las alarmas estarán expresadas como salidas de consola con información relevante. (Ver Figura 4.4)

```
Video de Día
-----
Distraction! (yaw)
Distraction! (yaw)
Distraction! (yaw)

PRIMER ALARMA ALCANZADA: Distracciones Leves
Indice de Somnolencia: 1
Somnolencia Controlada. Fuera de Riesgo.
```

Fig. 4.4: Imagen de disparo de la alarma 1 en detección de distracciones leves.

En cuanto a la información mostrada, se muestra en primer lugar el número de punto de control que corresponde según el momento del video, luego se muestra un mensaje basándose en el vector de distracciones que se fue cargando hasta ese momento, y finalmente muestra el índice de somnolencia calculado hasta esa instancia.

Junto con el número de alarma que se está ejecutando, se mostrará el mensaje Distracciones Leves, Distracciones Moderadas, Distracciones Peligrosas o Sin Distracciones si se da el caso. Así, el mensaje variará de la siguiente manera:

- Distracciones Leves: De 1 a 3 Distracciones contadas en el vector
- Distracciones Moderadas: De 4 a 5 Distracciones contadas en el vector.
- Distracciones Peligrosas: De 6 en adelante Distracciones contadas en el vector

Finalmente, luego de finalizado el video el sistema enumeró una serie de datos útiles los cuales sirvieron para evaluar el video y sacar conclusiones pertinentes (Ver Figura 4.5). Así, se enumeraron datos de relevancia tales como:

- Número total de frames
- Valor final del índice de somnolencia
- Cantidad de distracciones
- Segundo en el que se desarrolla cada distracción
- Segundos totales de duración del video
- Tipo de distracción que se detectó

```

Video: DL_T_M
Frames en total: 1854
Primera alarma en frame : 463
Segunda alarma en frame: 926
Tercera alarma en frame: 1389
-----
Frames no detectados: 14
Duración Video: 66 segundos
Indice de Distracciones: 71
Clasificación del Indice: Atención! Parpadeos excesivos
Total de Distracciones: 7
-----
Detalle de Distracciones:
-----
Distracción 1: Distraccion Parcial en segundo: 3
Distracción 2: Distraccion Parcial en segundo: 4
Distracción 3: Distraccion Parcial en segundo: 14
Distracción 4: Distraccion Parcial en segundo: 20
Distracción 5: Distraccion Parcial en segundo: 21
Distracción 6: Distraccion Parcial en segundo: 46
Distracción 7: Distraccion Parcial en segundo: 53
-----
Video Finalizado. Gracias por reproducir.

```

Fig.4.5: Pantalla final con información de relevancia

4.5 Resultados y discusión

Para ver los resultados de la implementación se analizó por un lado la base de datos creada para el estudio junto con todos los parámetros obtenidos en la investigación, y por otro lado se analizaron los mismos parámetros en un conjunto dedicado a pruebas.

4.5.1 Resultados sobre el conjunto inicial

Para evaluar el desempeño de la implementación, se realizó una comparación, entre los tiempos que se detectaron de forma empírica o manual, y los tiempos detectados por el sistema tal como se aprecia en la Tabla 4.3. Es necesario recordar que previamente junto a la grabación de los videos se desarrollaron para cada uno de ellos una ficha con la información relevante de cada escenario. Es así, que se compararon video a video los segundos de distracción detectados en cada instancia y así poder sacar una conclusión adecuada.

Se tomaron en cuenta las devoluciones del sistema en tiempo real, por lo que los números referidos a las distracciones ("Dist." en la tabla) hacen referencia a distracciones por el movimiento de la cabeza del conductor. También se analiza el índice de somnolencia en cada caso ("Somno." en la tabla) ya que varias de las distracciones tienen raíz en este fenómeno.

Video	Dist. 1 (ficha)	Dist. 1 (sistema)	Dist. 2 (ficha)	Dist. 2 (sistema)	Dist. 3 (ficha)	Dist. 3 (sistema)	Somno.
Leves (1 a 3 segundos)							
DL_M_M	3	3	2	1	1	2	23
DL_M_V	2	1	3	3	2	-	11
DL_T_M	4	2	2	2	3	4	38
DL_T_V	1	-	1	1	3	3	16
DL_N_M	2	1	2	-	1	1	30
DL_N_V	3	3	2	1	3	4	27
Moderadas (4 a 6 segundos)							
DM_M_M	4	4	3	5	5	4	29
DM_M_V	5	3	3	3	4	6	17
DM_T_M	3	4	4	4	6	4	50
DM_T_V	5	5	5	4	4	6	26
DM_N_M	4	5	5	4	6	3	27
DM_N_V	4	6	4	3	3	1	40
Peligrosas (mayores a 6 segundos)							
DP_M_M	6	6	7	5	8	7	90
DP_M_V	7	7	8	6	6	5	84
DP_T_M	7	5	7	7	6	7	34
DP_T_V	7	-	6	1	5	-	117
DP_N_M	8	2	8	-	7	-	110
DP_N_V	6	-	7	3	7	-	120

Tabla 4.3: Tabla de comparación entre las fichas correspondientes a los videos y las detecciones. Distracciones medidas en segundos y valor del índice de somnolencia

Por lo percibido en los resultados se pueden apreciar números satisfactorios en lo que respecta a las detecciones leves y moderadas. Las distracciones por el movimiento de la cabeza son detectadas de forma acertada, y los números del índice de somnolencia corresponden a números normales para esos grupos.

Sin embargo, cuando se trata de distracciones peligrosas debemos considerar que se producen por dos causas que consideramos de extremo peligro: las distracciones peligrosas por movimiento y el índice de somnolencia. Se han detectado muchos casos donde un sujeto con somnolencia no mueve su cabeza tanto como para que se detecte una distracción peligrosa por movimiento pero se considera peligrosa por el elevado grado de somnolencia. Puede observarse por ejemplo en el último video de la lista, donde prácticamente no se detectan distracciones por movimiento de la cabeza del conductor, sin embargo tiene un índice de somnolencia de 120, el cual equivale según lo explicado en la sección 4.3.1 a 48 segundos de somnolencia.

Se muestra a continuación un resultado adicional, donde se evalúa de forma general las detecciones por grupo donde se comparan el total de segundos detectados por el sistema, contra el total de segundos definidos en las fichas:

Tipo de Distracción	Segundos de las fichas	Segundos detectados	Porcentaje de acierto
Leves	40	32	80%
Moderadas	77	74	96%
Peligrosas	124	61	49%

Tabla 4.4: Tabla de comparación entre las fichas correspondientes a los videos y las detecciones. Tabla dividida según tipo de distracción

La Tabla 4.4 resume los datos presentados en la Tabla 4.3, de modo que en primer lugar se visualizan porcentajes de acierto muy positivos en lo que respecta a las distracciones leves y moderadas.

Por otro lado, en cuanto a las distracciones peligrosas se puede observar un porcentaje de acierto relativamente bajo. Esto se debe a que se contabilizaron los segundos en los que el conductor se distrajo en las fichas y las detecciones que el sistema hizo por movimientos de cabeza, discriminando los episodios de somnolencia, que como ya se mencionó anteriormente, fueron analizados a través del índice de somnolencia.

Para ver un análisis cualitativo de las distracciones, se muestra a continuación en la Tabla 4.5 un cuadro comparativo para ver cómo fueron detectadas las distracciones. Para la construcción del cuadro, se utilizaron los informes obtenidos en la sección 4.4.2 (Ver figura 4.6), donde están presentes los resultados de cada video. Así, se tomaron los videos de distracciones por grupo leves/moderadas/peligrosas, y se contabilizó cuantas eran del tipo parcial, y cuantas del tipo total, obteniendo así medidas en porcentajes.

```

Video: DL_T_M
Frames en total: 1854
Primera alarma en frame : 463
Segunda alarma en frame: 926
Tercera alarma en frame: 1389
-----
Frames no detectados: 14
Duración Video: 66 segundos
Indice de Distracciones: 71
Clasificación del Índice: Atención! Parpadeos excesivos
Total de Distracciones: 7
-----
Detalle de Distracciones:
-----
Distracción 1: Distraccion Parcial en segundo: 3
Distracción 2: Distraccion Parcial en segundo: 4
Distracción 3: Distraccion Parcial en segundo: 14
Distracción 4: Distraccion Parcial en segundo: 20
Distracción 5: Distraccion Parcial en segundo: 21
Distracción 6: Distraccion Parcial en segundo: 46
Distracción 7: Distraccion Parcial en segundo: 53
-----
Video Finalizado. Gracias por reproducir.

```

Fig.4.6: Ejemplo de informe final de cada video

Se debe recordar que a los conductores, solo se les aclaró el tipo de distracción que debían realizar, esto fue distracciones leves, moderadas o peligrosas y nunca se les dijo “como” distraerse, de modo que podían perder la atención en el camino de forma parcial o total.

Tipo de distracción	Leve	Moderada	Peligrosa
Pérdida total de visión del camino	17%	62%	84%
Pérdida parcial de visión del camino	83%	38%	16%

Tabla 4.5: Tabla de comparación entre los dos grupos de distracciones integrados.

Lo que es importante ver de este cuadro es la calidad de las distracciones, podemos ver que en las distracciones leves, predominaron distracciones parciales, donde el conductor se distrae muy poco de su ámbito de conducción, mientras que en las peligrosas, hay pérdidas totales de la visión en casi todos los videos, esto está relacionado con la somnolencia presente en los videos.

4.5.2 Resultados sobre conjunto de prueba

Una vez calibrado todos los parámetros en el conjunto inicial, se pasó a evaluar el sistema sobre un conjunto de pruebas. Este conjunto consistió en 5 videos con características similares al conjunto de entrenamiento (videos de misma duración, con 3 distracciones principales, filmadas desde el mismo lugar) y fueron analizados de la

misma forma en que se evaluaron los videos previos es decir contrastando con su ficha técnica.

El conjunto tiene pocos videos debido a que esta etapa de evaluación no fue considerada en etapas tempranas del proyecto y algunos de los elementos que se tenían cuando fueron filmados los videos iniciales, ya no se poseen (disponibilidad del automóvil y sujetos de prueba principalmente). Por ende se decidió filmar 5 videos con las siguientes características:

- Un video de distracciones leves/mañana/hombre.
- Un video de distracciones moderadas/ mañana /hombre.
- Un video de distracciones peligrosas/mañana/hombre (distracciones por movimiento de cabeza)
- Un video de distracciones peligrosas/mañana/hombre (distracciones por somnolencia)
- Un video de distracciones peligrosas/noche/hombre (distracciones por somnolencia)

Al igual que se realizó previamente, se muestra en la Tabla 4.6 los resultados que mostró el sistema en contraste con las fichas creadas para cada video:

Video	Dist. 1 (ficha)	Dist. 1 (sistema)	Dist. 2 (ficha)	Dist. 2 (sistema)	Dist. 3 (ficha)	Dist. 3 (sistema)	Somno.
Leves (1 a 3 segundos)							
DL_M_H	2	2	3	2	1	1	16
Moderadas (4 a 6 segundos)							
DM_M_H	4	3	5	4	4	4	30
Peligrosas (mayor a 6 segundos)							
DP_M_H	7	6	6	6	6	6	77
DP_M_H	6	-	7	1	6	1	121
DP_N_H	6	1	6	-	5	1	100

Tabla 4.6: Tabla de comparación entre las fichas correspondientes a los videos de prueba y las detecciones.

Al igual que en los análisis previos se repite el escenario en que se detectan con éxito las distracciones por movimiento de la cabeza del conductor en los escenarios que no implican somnolencia (los 3 primeros casos). Para los últimos dos, se pueden apreciar números altos en lo que respecta al índice de somnolencia (121 en el penúltimo video que equivalen a 48,4 segundos de somnolencia según la tabla 4.2 y 100 en el último video que equivale a 40 segundos de somnolencia).

Algo que se puede observar es que el índice de somnolencia es un poco menor en el último video con respecto al penúltimo, esto puede deberse a que se trató de un video de noche, y al ser un video con el que nunca se probaron parámetros el resultado pudo

variar. Sin embargo el número resultante sigue dentro de la franja de índices correspondientes a las distracciones peligrosas, por tanto se considera como un resultado satisfactorio.

Se deja a continuación una tabla con resultados generales según el tipo de distracción tal como se hizo en la sección anterior:

Tipo de Distracción	Segundos de las fichas	Segundos detectados	Porcentaje de acierto
Leves	6	5	83%
Moderadas	13	11	84%
Peligrosas	55	22	40%

Tabla 4.7: Tabla de comparación entre las fichas correspondientes a los videos de prueba y las detecciones. Tabla dividida según tipo de distracción

Similar a los videos del conjunto inicial, la tabla muestra resultados favorables en cuanto a los dos primeros tipos de distracciones. Algo que es necesario destacar, es que los números tanto de las fichas como de las detecciones son pequeños debido a la cantidad reducida de videos con los que se trabajó. En cuanto al tercer grupo, se ve un porcentaje de acierto nuevamente bajo debido a la presencia de episodios de somnolencia en los videos.

A continuación se pasan a mostrar las devoluciones del sistema en cuanto a la clasificación cualitativa de las distracciones, es decir, si fueron pérdidas totales de visión o perdidas parciales:

Tipo de distracción	Leve	Moderada	Peligrosa
Pérdida total de visión del camino	20%	60%	70%
Pérdida parcial de visión del camino	80%	40%	30%

Tabla 4.8: Tabla de comparación entre los dos grupos de distracciones integrados en videos de prueba.

Éstos números como ya se mencionó hacen referencia a los resultados que provienen de los informes al finalizar el video. Por ejemplo al haber un solo video de distracciones leves, se vio el informe del video y se observó que el 80% de las distracciones fueron de pérdida parcial del camino.

Esto se realizó con todos los videos y al igual que los videos de entrenamiento dieron números semejantes al verificar los tipos de distracciones.



Fig. 4.7: Capturas pertenecientes a los videos de prueba

Para una implementación futura se pretende buscar la detección de distracciones con mayores complejidades, refiriéndose así a situaciones en donde el conductor presente gafas oscuras, donde haya casi nula luminosidad, entre otras.

En los últimos capítulos se habló sobre metodologías para la detección y clasificación de distracciones y pérdidas de atención en personas dedicadas al ámbito de la conducción. Como conclusión y viendo la robustez del sistema se puede afirmar que el mismo podría ser implementado en otros vehículos sin ninguna restricción, además, cabe mencionar que el pre-proceso que se le hizo a las imágenes nocturnas fue acertado, por lo que serían datos que quedarían disponibles para futuros proyectos.

- Capítulo V –

5. Interacción con el usuario y componentes visuales

En esta etapa del proyecto, se le da énfasis a los aspectos relacionados con el comportamiento que presenta el sistema con respecto a la interacción con el usuario. Así, se plantea la utilización de distintas metodologías para que el sistema sea lo más amigable posible en términos de utilización.

Es menester mencionar que ésta instancia del proyecto fue destinada a un usuario orientado al desarrollo. Dejándose para una instancia futura, la orientación del proyecto para un usuario final de una aplicación.

Se trabaja además sobre la parte final del sistema, generando componentes tales como informes de conclusiones, con el fin de que el usuario pueda ver los resultados generados en el análisis.

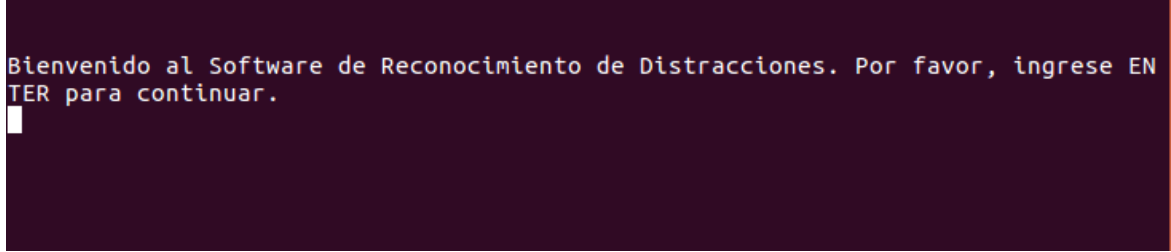
5.1 Menú principal y videos de entrada

5.1.1 Interacción con usuario

Una de las pautas que se tomaron en cuenta para realizar una experiencia de usuario, consiste en la utilización de una consola de comandos para la interfaz, ya que se hace más fácil de seguir a la hora de mostrar el funcionamiento del sistema a otro desarrollador. En un principio se pensó la utilización de interfaces gráficas, sin embargo por cuestiones de tiempos de entrega y velocidad de desarrollo, se decidió la utilización de la consola de comandos. [26]

Como modalidad de comunicación con el usuario, se elige una metodología de guía, con el fin de que el usuario pueda ingresar al sistema, y sin grandes confusiones pueda entender el funcionamiento del proceso a evaluar. Para lograr este efecto de guía, se implementaron funciones que dan tiempo al evaluador de leer todo el menú de interacción, como así también funcionalidades en las que el usuario debe presionar una tecla cualquiera para continuar (Ver Figura 5.1). Esto logra que cada usuario además

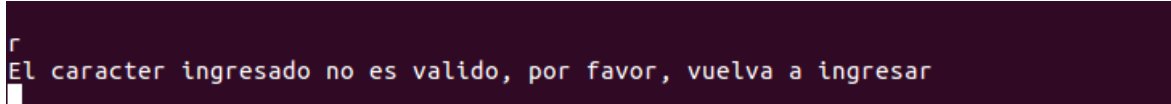
de poder analizar el programa de una forma directa, también pueda hacerlo a su debido tiempo.



```
Bienvenido al Software de Reconocimiento de Distracciones. Por favor, ingrese ENTER para continuar.  
█
```

Fig. 5.1: Pantalla de bienvenida.

Para evitar que el programa ingrese en situaciones no deseadas – como por ejemplo si el sistema pide el ingreso de un número y el usuario ingresa una letra – se opta por la utilización de varias estructuras de decisión switch, que combinadas con estructuras de iteración como While, permitan que el usuario pueda usar el sistema de forma correcta, sin ninguna ambigüedad (Ver Figura 5.2).



```
█  
El caracter ingresado no es valido, por favor, vuelva a ingresar
```

Fig. 5.2: Respuesta cuando el usuario ingresa una letra en lugar de un número.

5.1.2 Selección de video de entrada

En esta implementación para desarrollo y evaluación, el sistema utiliza videos alojados en la memoria del dispositivo, y se posterga el uso de videos en tiempo real para la versión de usuario final. Para la elección del video a evaluar, el sistema pedirá al usuario que ingrese según cierto número, el video de preferencia. Cada uno de estos videos, deberá estar contenido en la carpeta raíz del programa, y con su nombre correspondiente. Por ejemplo, si se tratase de un video de distracciones leves, durante la tarde, ejecutado por una mujer, el video deberá tener el título de DL_T_M.mp4. Según el número que elija el usuario, habrá una cadena de caracteres parecida a la mencionada que será la que dispare el video escogido (Ver Figura 5.3).

```

A continuación, ingrese el video a evaluar en base a los siguientes items:
1) Distracción leve / Mañana / Varón
2) Distracción leve / Mañana / Mujer
3) Distracción leve / Tarde / Varón
4) Distracción leve / Tarde / Mujer
5) Distracción leve / Noche / Varón
6) Distracción leve / Noche / Mujer

7) Distracción Moderada / Mañana / Varón
8) Distracción Moderada / Mañana / Mujer
9) Distracción Moderada / Tarde / Varón
10) Distracción Moderada / Tarde / Mujer
11) Distracción Moderada / Noche / Varón
12) Distracción Moderada / Noche / Mujer

13) Distracción Peligrosas / Mañana / Varón
14) Distracción Peligrosas / Mañana / Mujer
15) Distracción Peligrosas / Tarde / Varón
16) Distracción Peligrosas / Tarde / Mujer
17) Distracción Peligrosas / Noche / Varón
18) Distracción Peligrosas / Noche / Mujer

19) Reproduzca un archivo a elección

```

Fig.5.3: Listado de videos preconfigurados más la opción de carga de un video particular.

Además de los 18 videos constituidos previamente, el software da la opción de introducir un video propio, distinto a los predeterminados. Para utilizar dicha alternativa, basta con elegir la última opción del menú de videos, e introducir el nombre del video a reproducir, creando así una cadena de caracteres que tendrá como finalidad reproducir el archivo con ese nombre. Para ejecutar esta última alternativa se debe tener en cuenta el formato del video, ya que el entorno está configurado para reproducir videos del tipo mp4, avi y mpv.

5.2 Elementos visuales adicionales

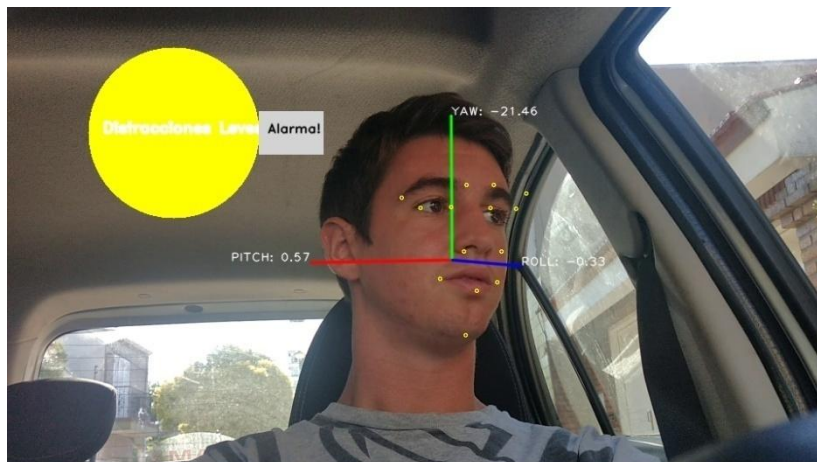
Para colaborar con la oferta visual que brinda el sistema, se propone la señalización de las diferentes alarmas dependiendo si las mismas son leves, moderadas o peligrosas. Cada 15 segundos, donde se evalúan las distracciones de cada video, se presenta un señalador en forma de círculo (Ver Figura 5.4), si el video viene presentando distracciones leves (círculo amarillo), moderadas (círculo naranja) o peligrosas (círculo rojo). Como también se presenta un señalador celeste si el mismo presenta ausencia de distracciones. [27]

En el ámbito de las señalizaciones, también se habla de señalizaciones auditivas, es por eso que ante la aparición de distracciones en exceso – de más de 3 segundos – se reproducen sonidos predeterminados. Para reflejar éste concepto, se optó por la

colocación de una imagen de un pequeño cuadrado gris (simbolizando un megáfono) en una parte de la pantalla, con el fin de demostrar que en una hipotética implementación física, el auto emitiría un pequeño ruido de advertencia.



(I)



(II)

Fig. 5.4: (I) Señalización visual de la aparición de Distracción Peligrosa. (II) Señalización visual de Distracción Leve con señal sonora.

5.3 Elaboración de informes

5.3.1 Creación de archivos

Para obtener los resultados pertenecientes a un video, el usuario tendrá la posibilidad de elegir la elaboración de un informe en forma de archivo de texto o en la misma interfaz. Esto es conseguido a través de un diálogo de confirmación, en donde el usuario deberá colocar "Y" si desea generar un archivo de texto, o "N" si desea ver los resultados sobre la consola (Ver Figura 5.5).

```
Desea crear un informe con los resultados? [Y/N]
```

(I)

```
INFORME DE DISTRACCIONES
Video DL_T_M.mp4

Frames en total: 1854
Primera alarma en el segundo : 16
Segunda alarma en el segundo: 33
Tercera alarma en el segundo: 49
-----
Frames no detectados: 14
Duración Video: 66 segundos
Indice de Somnolencia: 71
Total de Distracciones: 7
-----
Detalle de Distracciones:
-----
Distracción 1: Distraccion Parcial en segundo: 3
Distracción 2: Distraccion Parcial en segundo: 4
Distracción 3: Distraccion Parcial en segundo: 14
Distracción 4: Distraccion Parcial en segundo: 20
Distracción 5: Distraccion Parcial en segundo: 21
Distracción 6: Distraccion Parcial en segundo: 46
Distracción 7: Distraccion Parcial en segundo: 53
-----
Video Finalizado. Gracias por reproducir.
```

(II)

```

^DL_T_M.mp4 Informe (-/Facultad/ProyectoFinal/Opencv-headpose/Informes Distraccion/2018/10/17) - gedit
Abrir  [F]
INFORME DE DISTRACCIONES
Video DL_T_M.mp4

Frames en total: 1854
Primera alarma en segundo : 16
Segunda alarma en segundo: 33
Tercera alarma en segundo: 49
-----
Frames no detectados: 14
Duración Video: 66 segundos
Índice de Distracciones: 71
Total de Distracciones: 7
-----
Detalle de Distracciones:
-----
Distracción 1: Distraccion Parcial en segundo: 3
Distracción 2: Distraccion Parcial en segundo: 4
Distracción 3: Distraccion Parcial en segundo: 14

Hasta ahora se trató de distracciones LEVES

Distracción 4: Distraccion Parcial en segundo: 20
Distracción 5: Distraccion Parcial en segundo: 21

Hasta ahora se trató de distracciones LEVES

Distracción 6: Distracción Parcial en segundo: 46
Hasta ahora se trató de distracciones LEVES

Distracción 7: Distraccion Parcial en segundo: 53
Hasta ahora se trató de distracciones LEVES
-----
Video Finalizado. Gracias por reproducir.

```

(III)

Fig. 5.5: (I) Consulta al usuario sobre la creación del informe final. (II) Informe creado sobre la pantalla de comandos. (III) Informe creado en forma de archivo de texto, en la carpeta correspondiente.

5.3.2 Creación de carpetas

En caso de que el usuario haya optado por la generación de un archivo, el sistema consultará la fecha. Esto lo realiza con el fin de crear una jerarquía de carpetas, en la cual se puede mantener un orden entre los distintos archivos generados. El programa descompondrá la misma en sus distintos componentes, tanto año, mes y día.

El sistema crea un directorio llamado "Informes de Distracciones" el cual será creado en la misma dirección que se encuentra albergado el sistema. Esto lo realiza cada vez que se ejecuta el programa, si es la primera vez que el usuario ejecuta el mismo, el sistema creará el directorio. Cabe mencionar que el sistema enviará un mensaje de confirmación, haya creado o no el directorio, como se aprecia en la Figura 5.6. [27]

```
Desea crear un informe con los resultados? [Y/N]
y
La carpeta Informes ya fue creada
La carpeta Anio ya fue creada
La carpeta Mes ya fue creada
Carpeta Dia Creada !
```

Fig. 5.6: Notificaciones sobre las carpetas creadas en forma previa, más la creación de la carpeta “Día”.

Así como el sistema creó la carpeta raíz “Informes de Distracciones”, creará también dentro de ella la del año correspondiente al ordenador que ejecuta el sistema. Dentro de dicha carpeta creará la del mes, y más tarde la del día, donde en el interior de ésta última estarán introducidos los informes que genera el usuario. Finalmente, como comentario final se hace necesario mencionar que cada informe tendrá como nombre de archivo, el nombre del video referente que se analizó, más la palabra “Informe”. Donde sí por ejemplo se ha analizado el archivo “DL_M_V.mp4” – refiriéndose a un video de distracciones leves, de mañana, varón – el nombre del archivo de informe será “DL_M_V.mp4 Informe”.

5.4 Discusión

En el ámbito de la comunicación sistema-usuario, se evaluó que al tratarse de un trabajo de investigación y desarrollo, la modalidad de consola elegida resultó adecuada ya que los usuarios no pierden la atención con diferentes opciones como podría presentarse en el caso de una interfaz gráfica. Sino que a través de la consola de comandos, el sistema guía al usuario, hasta un resultado final que a su vez está presentado de una forma sencilla como son los informes distribuidos en sus respectivas carpetas.

Además se da la opción de ver los resultados por consola, por lo que le permite al usuario la posibilidad de plantear sus conclusiones en el momento. [21]

- Capítulo VI -

6. Conclusiones y desarrollos futuros

A lo largo de todo el proceso tanto de captura de imágenes como de desarrollo se pudieron apreciar diferentes resultados y diferentes variables que pudieron ser explotadas para obtener las conclusiones que se fueron viendo a lo largo de los capítulos.

Para la detección y análisis de pérdidas de la atención en la conducción, se procedió a la grabación de una serie de videos con características similares en cuanto a metodología de grabación, duración, y formato, los cuales sirvieron como entrada para la evaluación de diferentes mecánicas en lo que respecta a la detección de distracciones de distinto tipo.

Ante la presencia de factores peculiares como presencia de oscuridad en los videos, se procedió a un preprocesamiento de los mismos, mediante el análisis de los histogramas de la escena y la posterior aplicación del método de Corrección Gamma. Todo esto con resultados sumamente aceptables, y tomando una región de interés dentro de la escena para evitar analizar elementos poco relevantes.

Se plantearon dos grupos de distracciones. Por un lado, se analizaron distracciones catalogadas por severidad, es decir, distracciones leves, moderadas, y peligrosas, y por otro lado, se clasificaron distracciones según la pérdida de visión del camino, clasificadas en distracciones con pérdidas parciales de visión en el camino, o distracciones con pérdidas totales de visión en el camino.

Para el primer grupo, se tomó en cuenta la duración de las distracciones y mediante varias pruebas se llegó a ciertos criterios para clasificarlas. En cuanto al segundo, para las distracciones con pérdidas de la atención total, se tomó en cuenta los momentos de ausencia de detección de rostro por parte de la librería, es decir, cuando el conductor deja de mirar el camino en su totalidad, mientras que para las distracciones parciales, se calcularon ángulos de Euler con el fin de detectar cuando el conductor desvía su mirada de la ruta pero solo en una pequeña proporción.

Para situaciones como la somnolencia, se utilizó un modelo matemático el cual relaciona las distancias entre los distintos puntos de los ojos con el fin de realizar una cuantificación de los parpadeos. Es así que contando los mismos se pudo ir sacando

un índice el cual va informando sobre el estado de somnolencia del conductor, así, fue catalogado en el presente escrito como “Índice de somnolencia”. Para cuantificar este índice se hizo una equivalencia en segundos de modo de poder ver qué porcentaje del video corresponde a un comportamiento somnoliento.

Cada uno de los análisis realizados a su vez, pueden ser exportados a archivos de texto, en una carpeta raíz la cual tendrá el nombre del día y del mes en el cual fue realizada la prueba. Con todos los datos pertinentes para la correcta evaluación del video.

Toda la metodología y las herramientas implementadas, fueron desarrolladas también con el fin de brindar una aplicación para desarrolladores con el propósito de implementar sus propias soluciones y probar distintos métodos para mejorar las detecciones propuestas.

En un futuro se buscará mejorar la robustez del presente sistema, con el fin de que se expanda a nuevas situaciones las cuales no fueron tomadas en cuenta al momento de realizar el proyecto. Situaciones como conductores con gafas de Sol, autos descapotables, días lluviosos, nublados, luces del alumbrado público, luces de otros vehículos, Sol de frente, de costado, etc. son algunas de las situaciones las cuales no fueron tomadas en cuenta a la hora de desarrollar metodologías de detección.

Otra de las vías hacia el futuro podría ser la implementación de los métodos ya en una plataforma móvil, ya pensando en un usuario final, se hablaría de una aplicación la cual tendría activada su cámara, sea delantera o trasera, y mediante un soporte, colocar el teléfono en el automóvil, para que éste realice las detecciones y envíe alarmas cuando sea pertinente.

Referencias

- [1] [2] Diario Clarín: "Argentina, a la cabeza de un ranking de muertes en el tránsito". URL accedida el 13 de abril de 2017:
https://www.clarin.com/sociedad/itf-ocde-siniestros_viales-estadistica_internacional_0_ryfsoOFv7g.html
- [3] Centro de Experimentación, Investigación y Seguridad Vial (CESVI) - Estadísticas sobre Seguridad Vial: Accidentes relevados por causa, accidentes relevados por tipo de vehículo, fallas humanas más comunes. URL accedida el 24 de abril de 2017:
<https://www.cesvi.com.ar/SeguridadVial/Estadisticas/SeguridadEstadisticas.aspx>
- [4] Ministerio del Interior. Subdirección general de intervención y políticas viales. Unidad de intervención educativa. "Distracciones al volante". URL accedida el 13 de abril de 2017:
http://www.dgt.es/PEVI/documentos/catalogo_recursos/didacticos/did_adultas/Distracciones_al_volante.pdf
- [5] Actualidad Motor: "Bosch desarrolló Sistema de Detección del Sueño". URL accedida el 13 de abril de 2017:
<https://www.actualidadmotor.com/bosch-desarrollo-sistema-de-deteccion-del-sueno/>
- [6] Librería Dlib. URL accedida el 13/03/2018: <http://Dlib.net/>
- [7] Facial landmarks using Dlib, opencv, and python. URL accedida el 13/3/18:
<http://www.pyimagesearch.com/2017/04/03/facial-landmarks-Dlib-opencv-python/>
- [8] [9][10] Bases de datos consultadas. URLs accedida el 13/3/2018:
- <http://www.robesafe.com/personal/jnuevo/Datasets.html>
 - <http://adas.cvc.uab.es/elektra/enigma-portfolio/cvc11-drivface-dataset/>
 - <https://www.kaggle.com/c/state-farm-distracted-driver-detection/data>
- [11] Sensor Omnivision (OV16860) URL accedida el 13/3/18:
<http://www.ovt.com/sensors/OV16860>
- [12] OpenCV: Clasificadores en cascada LBP y Haar. URL accedida el 18/3/20:
https://docs.opencv.org/master/dc/d88/tutorial_traincascade.html

- [13] Viola P, Jones M.J. (2004) International Journal of Computer Vision. Robust Real-Time Face Detection. (Pages 57-89)
- [14] WANG Y., Haizhou A., Bo W., Chang H. Dept. of Computer Science and Technology, Tsinghua University, State Key Laboratory of Intelligent Technology and Systems. Real Time Facial Expression Recognition with Adaboost
- [15] Osimani C. (2014) Centro de Investigación Aplicada y Desarrollo en Informática y Telecomunicaciones. Universidad de Blas Pascal. *Análisis y procesamiento de imágenes para la detección del contorno labial en pacientes de odontología* (Page 3)
- [16] J. Soler del Pozo (Junio 2009) Universidad de Murcia. Facultad de Informática. Ingeniería superior informática. Proyecto final de carrera. *Detección de matrículas mediante características binarias locales (LBP)* (Pages 26-27)
- [17] Histogramas y estimación de rostro en metodologías LBP: <http://campus.unca.edu.mx/expotics/reconocimiento-de-rostros>
- [18] Facial landmarks using Dlib, opencv, and python. URL accedida el 13/3/18: <http://www.pyimagesearch.com/2017/04/03/facial-landmarks-Dlib-opencv-python/>
- [19] Face detection functions – Opencv. URL Accedida el 27/3/2020: https://docs.opencv.org/3.4.9/db/d7c/group__face.html
- [20] Estimar postura 3D – Opencv y Dlib. URL Accedida el 11/07/18: <http://acodigo.blogspot.com/2017/11/estimar-postura-3d.html>
- [21] Arfken G. (1985) Mathematical Methods for Physicists, 3rd. Edition. Orlando, FL: Academic Press. Euler Angles. (Pages 198-200)
- [22] Unidad de medida Inercial. Algoritmo de Estimación e Implementación Software Desarrollo teórico. URL Accedida el 27/3/20: http://bibing.us.es/proyectos/abreproy/11253/fichero/memoria%252F02_capitulo2.pdf
- [23] (1) Gonzalez C.R, Woods E.R. (2008) Digital Image Processing Third Edition. Gamma Transformation (Pages 132-137)
- [24] (2) Eye blink detection with OpenCV, Python, and Dlib. URL Accedida el 11/08/18: <https://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-Dlib/>

- [25] M. Gomez Montiel (2015) Estudio de las características del parpadeo, y su relación con los movimientos sacádicos, en distintas condiciones controladas de lectura (Pages 6-7)
- [26] Stroustrup B. (2002) The C++ Programming Language, Pearsons Education. (Pages 899 - 906)
- [27] Gonzalez C.R, Woods E.R. (2008) Digital Image Processing Third Edition. Drawing Functions (Pages 132-137)
- [28] Shotts Jr. W., (2012) The Linux Command Line: A complete introduction. No Starch Press. Making Directory Commands. (Pages 300 - 308)
- [29] Sitio web Google Drive con ejemplos de los videos y su funcionalidad. URL Accedida el 9/11/2018:
https://drive.google.com/drive/folders/105Wt_bxBa6YhwFp8DDIYUNZCd6vkID1R?usp=sharing

sinc(r) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)

G. Farré, E. M. Albormoz & C. E. Martínez; "Desarrollo de un sistema de alerta de pérdida de atención en la conducción basado en visión computacional (Undergraduate project)"
Facultad de Ingeniería y Ciencias Hídricas - Universidad Nacional del Litoral, 2020.

Plan del proyecto y manual de usuario



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas

PROYECTO FINAL DE CARRERA
INGENIERÍA INFORMÁTICA

**Desarrollo de un sistema de detección de pérdidas de atención en la
conducción basado en visión computacional**

Alumno: Farré, Gonzalo

Director: Martínez, César

Co-Director: Albornoz, Enrique Marcelo

Resumen

Se está viviendo en un mundo tanto consumista como acelerado, y mientras la tecnología está acortando distancias continentales a la simple idea de un click, las distancias físicas siguen siendo las mismas. La sociedad se moviliza a través de vehículos motorizados cada vez más modernos y más tecnológicos día tras día, sin embargo a pesar de que los avances sean cada vez más complejos, la seguridad vial sigue siendo uno de las principales causas de preocupación en lo que respecta a nuestro andar diario.

Las distracciones al conducir son una de las principales causas de accidentes de los últimos tiempos, y a pesar de que éstos hechos se repitan una y otra vez, pocas son las cosas que se llevan a cabo para prevenirlos.

Es así que en el presente proyecto se planteará una alternativa para la detección y prevención de pérdidas de la atención al conducir mediante la ejecución de distintos tipos de alarmas basándose en la mirada del conductor y la dinámica temporal de la misma mediante la utilización de técnicas de procesamiento digital de imágenes.

Palabras clave: Distracción, alarmas, seguridad vial, dinámica temporal, procesamiento digital de imágenes.

Justificación

La utilización de automóviles para el traslado es un hecho totalmente habitual en la vida de hoy en día, hecho que no era así hace 30 años atrás. Sin embargo, a pesar de los avances tecnológicos que pudieron llegar a darse en el campo automovilístico, la seguridad vial sigue siendo una de las principales preocupaciones que sigue afectando el hecho de vivir en sociedad.

Hablando exclusivamente de nuestro país, Argentina encabeza una lista de 38 países en términos de muertes por causas viales, donde se habla de más de 5000 casos (**7.268 en 2016 [1]**), con una tasa de 12,3 muertes por año por cada 100.000 habitantes. La tasa más pequeña la tienen países como Suecia, con una tasa de 2,7 fallecimientos por año cada 100.000 habitantes [2]

Es importante destacar ahora una serie de datos que serán de utilidad para entender un poco más el contexto en el que se encuentra el proyecto y analizar la

problemática que se nos presenta. Así, de acuerdo con las estadísticas de CESVI (**Centro de Experimentación, Investigación y Seguridad Vial**) sobre la base del análisis de 4000 accidentes de tránsito ocurridos en los últimos 11 años, la mitad de los accidentes fueron protagonizados por autos. Lo siguen los camiones (23%), pick-ups (9%), utilitarios y 4x4 (8%), micros (3%), motocicletas (3%) y otros (4%). [3]

A su vez, es importante destacar las principales causas de dichos accidentes, siendo 90% de los choques debidos al factor humano, 5% al vehículo y 5% consecuencia de las malas condiciones de caminos y el clima. Ahora, viendo más en foco la principal causa de incidentes, se mencionan las principales falencias humanas que se registraron en dichos escenarios, donde según la entidad mencionada, 41% de las fallas humanas que llevan a los accidentes es la invasión de carril. Siguen: la distracción (19%), la velocidad inadecuada (16%), las maniobras abruptas (10%), distancia inadecuada (6%), prioridad de paso no respetada (4%), cansancio (4%). Obviamente, estarían incluidas dentro de estas fallas las causadas por la ingesta de alcohol. [ibid]

Las distracciones, razones principales para la pérdida de atención, son definidas según la Real Academia Española, como “cosa que atrae la atención apartándola de aquello a lo que está aplicada...”. Con esta definición, se puede catalogar una serie de distracciones comunes las cuales podrían decirse que eventualmente participan en los accidentes mencionados anteriormente, entre ellas se mencionan: utilización del teléfono celular, uso del cigarrillo, maquillaje, búsqueda de objetos, etc. [4]

Para ver como se le hace frente a éstos accidentes en la actualidad, juega un papel importante Bosch, empresa alemana vinculada a la construcción y distribución de partes y sistemas de la industria automovilística, la cual ha desarrollado un sistema de detección de fatiga aplicable a casi cualquier tipo de auto, el cual se encarga de utilizar la información, ya sea del sistema eléctrico de dirección asistida o del sistema de control de estabilidad, a partir de la cual evalúa las acciones del conductor y busca signos de cansancio y sueño.

Por medio de la utilización de un sensor puede determinar los hábitos de conducción del conductor, esto significa que ya no será necesario utilizar sensores de radar altamente costosos que se fundamentan en la utilización de un sistema de

cámara y video. Así, este desarrollo reduce significativamente el costo, lo que resulta altamente rentable para los coches de gamas más bajas. [5]

Lo que se pretende con este proyecto será, teniendo en cuenta las estadísticas mencionadas, construir un sistema que no sólo se encargue de las fatigas del conductor, sino que se ocupe de las distracciones del mismo (siendo las distracciones la causa del 19% de las falencias humanas contra sólo un 4% que aporta el cansancio y la fatiga). Se atacarán las distracciones mencionadas midiendo la mirada del conductor y el tiempo que ésta se desvía de la conducción del automóvil. Detectando además, la somnolencia de la cual se ocupa Bosch.

Para ocuparse de las pérdidas de atención, lo que se pretende es la utilización de una librería llamada Dlib utilizada sobretodo en el ámbito del procesamiento digital de imágenes faciales. Lo que ésta librería permite, es mediante la utilización de marcadores o "landmarks", detectar una gran cantidad de puntos de interés a lo largo del rostro para luego evaluar el comportamiento de estos puntos al ejercer alguna expresión o gesto. Así, se pretende darle gran importancia a los marcadores correspondientes a los ojos, ya que mediante la detección de los mismos se irá registrando la mirada del conductor, y así se definirá si éste pierde la atención al conducir o no.

Para la evaluación de la dinámica temporal lo que se busca es ir evaluando la detección de los ojos mediante los marcadores mencionados, y para el seguimiento de la mirada lo que se llevará a cabo es la utilización de un vector 3D normal al plano de la mirada del conductor. Todo esto llevado a cabo en una interfaz gráfica basándose en una base de datos compuesta por videos de prueba.

El sistema irá evaluando a través de videos la dinámica temporal de la mirada del conductor, brindando diferentes tipos de "Alarmas" de acuerdo a la cantidad de tiempo que éste pierde su atención al conducir, para esto se irá evaluando en los videos la cantidad de tiempo que el sistema deja de detectar la mirada del conductor y en base al tiempo detectado ejercerá un tipo de alarma para concientizar al sujeto. Éstas alarmas serán catalogadas como "Alarma tipo 1", "Alarma tipo 2" y "Alarma tipo 3", donde las mismas irán difiriendo entre sí en base a la cantidad de tiempo que el conductor pierde la atención. Refiriéndose a Alarma tipo 3" a una situación donde por

ejemplo se pierde la atención una gran cantidad de tiempo y podría referirse a un caso de somnolencia.

Como aporte al proyecto, se hace mención a otros métodos en el campo de interés como son los clasificadores en cascada, o también llamados "cascadas de Haar", los cuales mediante la utilización de ciertas características que se hallan en la imagen se irán proponiendo distintas "detecciones" y a medida que la imagen avanza en la cascada de clasificadores se irán descartando y refinando en el análisis, hasta llegar a una sola detección referida al objeto de interés. [6]

Con la realización de éste proyecto, se pretende de alguna manera crear conciencia de la existencia estos accidentes, mediante la evaluación de los videos planteados de modo que el conductor esté consciente de las conductas peligrosas al volante y trate de evitarlas en el futuro. También está en el campo de interés que la interfaz gráfica sirva a entidades gubernamentales para la enseñanza de la buena conducción y prevención de accidentes, como también se buscaría que organismos tales como escuelas o institutos utilicen el sistema a desarrollar con el fin de inculcar hábitos de la correcta conducción vial en las personas desde una edad temprana. Éste se podría decir que será el impacto social que podría provocar la implementación del proyecto, junto con un movimiento de concientización por parte de toda la comunidad de conductores.

Objetivos

Objetivo General

Desarrollar un sistema basado en visión computacional para detección de pérdida de atención en la conducción, mediante la evaluación de la dinámica temporal de la mirada del conductor.

Objetivos Específicos

- Relevar el estado del arte y requerimientos.
- Diseñar y desarrollar los bloques correspondientes al sistema.
- Formular una base de datos de videos con la cual se trabajará a lo largo del proyecto.
- Experimentar pruebas exhaustiva con los correspondientes videos.

- Asignar alarmas a cada una de las distracciones, haciendo a su vez una jerarquía de las mismas de acuerdo a la severidad del comportamiento.
- Realizar informes periódicos con el fin de ir documentando el avance.
- Desarrollar de un documento final junto con conclusiones y resultados.

Alcance y Limitaciones

El proyecto pretende el análisis de una base de datos de video con una cámara situada en un lugar estratégico del vehículo. Así, se puede decir que podría ser aplicado a todo vehículo de 2 ejes o más, sea auto, colectivo, camión, etc.

Además, se supondrá el mismo formato para todos los videos, es decir se ubicará la cámara siempre en el mismo lugar y con los mismos estereotipos de vehículos, es decir, vehículos convencionales con volante a la izquierda, con visión trasera, de circulación cotidiana en la ciudad. Por lo que se excluirá por ejemplo, situaciones donde el volante se encuentre a la derecha, situaciones con conductores utilizando anteojos de Sol, etc.

Como limitación, se puede decir también que se pretende la construcción de un sistema desarrollado con c++ para uso exclusivo de PC, con las correspondientes librerías a utilizar para el procesamiento digital de imágenes, y queda exento el escenario de desarrollo de un sistema embebido para integrar a la computadora de un vehículo el cual podría ser una extensión a futuro del presente proyecto.

Metodología

Para el alcance efectivo de los objetivos planteados en el proyecto, se implementará un modelo de proceso incremental, el cual aplicará secuencias lineales en forma escalonada a medida que avanza el calendario de actividades. Las secuencias mencionadas tendrán como principal meta la producción de entregables mediante los llamados "incrementos". Estos serán de naturaleza iterativa y convergen con mucha rapidez a versiones funcionales del software [7].

La razón que conlleva a la elección de esta metodología, se debe a que el proyecto está distribuido en tres grandes bloques los cuales son a su vez relativamente

dependientes entre sí. Se puede entonces decir que el proyecto tiene una estructura modular, donde la salida de cada uno de éstos módulos, serviría de entrada para la etapa siguiente. Así, haciendo mención a los distintos módulos, se menciona que en el ámbito del sistema de detección de pérdida de la atención, antes de la asignación de los distintos tipos de alarmas ante los distintos escenarios de distracción, el sistema debe ser capaz de detectar esa pérdida de atención al conducir, y para detectar dicho fenómeno, el sistema deberá tener a su disposición una correspondiente base de datos con la cual trabajar. Se pueden ver entonces las distintas fases del proyecto las cuales regirán la mayor parte del desarrollo.

Incremento 1: En la instancia de inicio lo que se pretende antes que nada es la obtención de una base de datos de video con la cual trabajar a lo largo del proyecto. Para esto lo que se realizará serán pruebas exhaustivas junto con la utilización de una cámara estratégicamente posicionada en el automóvil con el fin de realizar la captura de video correspondiente. Para llevar a cabo dicho incremento se consideraron las siguientes etapas:

- **Posicionamiento de la cámara en el automóvil:** Consistirá en la realización exhaustiva de pruebas para determinar la posición correcta que ocupará la cámara. Teniendo en cuenta los principales criterios para el desarrollo del proyecto, es decir que permita ver correctamente los ojos del conductor, manos del mismo, etc.
- **Realización de recorridos y captura de video:** Una vez instalado el dispositivo, se pretende la recolección de alrededor de 20 o 30 videos con diferentes escenarios de pérdida de atención en los mismos. Se pretenden videos de corta duración (un minuto o dos) en los que estén presentes los episodios de distracción más frecuentes como son utilización del teléfono celular, uso de cigarrillo, maquillaje, etc. y algunos casos de somnolencia
- **Documentación e informe del Incremento:** Finalmente se pretende la conformación de un informe con las conclusiones y los detalles de los distintos videos, clasificándolos por tipo de distracción, duración, alarma correspondiente, etc.

Incremento 2: En ésta instancia lo que se pretende es mediante la utilización de la librería Dlib [8] encontrar en los distintos videos las distintas pérdidas de atención que luego en el futuro generarán distintas alarmas. El enfoque se basa en utilizar la librería mencionada la cual se especializa en el reconocimiento de caras [9] [10] mediante la utilización de marcadores en los puntos de interés. Para luego utilizar estos puntos y realizar diferentes métricas para medir la falta de atención del conductor en base a los marcadores de los ojos. En esta instancia se pretende concretar las siguientes etapas:

- Estudiar conceptos relacionados con características faciales y la librería Dlib: Consiste en leer bibliografía sobre el campo de reconocimiento facial y estudiar las distintas funciones que nos brinda la librería para el marcado de puntos de interés.
- Software de reconocimiento de ojos: En base a videos correspondientes a la base de datos, utilizaremos la librería Dlib para fijar puntos de interés en los ojos con el fin de que sean los disparadores de las alarmas en caso de perder la atención al conducir. Se planea utilizar alrededor de 6 marcadores en cada ojo para el análisis junto con la utilización de un vector normal al punto de la mirada. [11]
- Documentación e informe del Incremento: Finalmente se pretende la conformación de un informe con las conclusiones y los detalles pertinentes a los desarrollos mencionados.

Incremento 3: En este punto, lo que se busca será a partir del reconocimiento anteriormente logrado, asignar a cada una de las distracciones un tipo de alarma, al basarse en el reconocimiento exclusivamente de los ojos, lo que se busca es emitir las alarmas en función de la cantidad de tiempo que el conductor se distrajo. Para esto se deben concretar las siguientes etapas:

- Métodos para la dinámica temporal: Desarrollo de métodos para la obtención de los momentos en los cuales se provocan las distracciones de interés dentro de los videos de la base de datos.

- Desarrollo de un método de alarmas, vinculada a las detecciones y a la dinámica temporal: Se buscará vincular las detecciones y sus respectivos momentos de aparición, con una alarma pertinente al comportamiento detectado, distinguiendo entre "alarma tipo 1", "alarma tipo 2" u "alarma tipo 3", haciendo mención de que los números asociados a las alarmas simbolizan su severidad. Siendo la "alarma tipo 3" la alarma más severa, disparada por un comportamiento tal como sería la somnolencia al conducir, la "alarma tipo 2" sería una alarma ya de prevención, como vendría a ser la utilización de un teléfono celular, y finalmente la "alarma tipo 1" será considerada como una alarma vinculada a un comportamiento relativamente inofensivo, como es la utilización de maquillaje, un cigarro, etc. Todas éstas serán disparadas no por el reconocimiento del objeto en sí, sino por la cantidad de tiempo que el conductor está perdiendo la atención. La "alarma tipo 1" en éste caso será activada por pérdidas de atención de pocos segundos, la "alarma tipo 2" por pérdidas mayores, como sería el caso de utilizar un teléfono, y la "alarma tipo 3" implica ya el hecho de la aparición de somnolencia por lo que será la alarma más importante. Éstas alarmas al ser un programa de C++ podría ser considerada una simple cadena de caracteres que se muestre en pantalla cuando ocurra un episodio deseado. Se deberán realizar todas las pruebas que se requieran.
- Documentación e informe del Incremento: Finalmente se pretende la conformación de un informe con las conclusiones y los detalles pertinentes a los desarrollos mencionados.

Incremento 4: Finalmente en éste último punto, se pretende la formación de una interfaz amigable, para que el usuario pueda fácilmente ver los videos de la base de datos, e ir viendo los distintos tipos de alertas que el sistema reconoce. Para finalizar esta última instancia se pretende:

- Formación de una interfaz gráfica: Interfaz gráfica con la posibilidad de ir seleccionando el video de interés y ver la funcionalidad del sistema en su conjunto.

- Redacción de Manual de Usuario: Redacción de un pequeño escrito con el fin de explicar las principales funcionalidades del sistema.
- Documentación e informe del Incremento: Finalmente se pretende la conformación de un informe con las conclusiones y los detalles pertinentes a los desarrollos mencionados.

Cada incremento estará vinculado con un entregable, los cuales serán corregidos por los correspondientes Directores del proyecto, para así validarlos y permitir así el correcto avance en el proyecto.

Entregables:

- **Incremento 1:** Como entregable se considerará a la base de datos de videos, la cual estará compuesta de 20 o 30 videos de duración no tan prolongada (1 minuto o 2), todos filmados desde el mismo ángulo y mostrando distintos tipos de pérdidas de atención. A su vez también habrá videos sin ningún tipo de distracción para evaluar la respuesta del software.
- **Incremento 2:** Se pretende entregar un software que a partir de los videos pertenecientes a la base de datos, reconozca el episodio de interés, basado en la utilización de la biblioteca Dlib.
- **Incremento 3:** Se mostrará una actualización del software utilizado de entregable, sólo que ahora se catalogarán las pérdidas de atención según la jerarquía de alarmas mencionadas ("alarma 1", "alarma 2", "alarma 3") junto con la dinámica temporal analizada.
- **Incremento 4:** El entregable corresponderá a una interfaz amigable con funcionalidades para ir pasando los distintos videos de la base de datos. Y en ella ir viendo los videos para analizar los resultados.

Plan de Tareas

1. Incremento 1 (60 horas)

1.1 Posicionamiento de cámara (10 horas)

1.1.1 Análisis de videos de prueba con distintas ubicaciones (5 horas)

1.1.2 Pruebas exhaustivas con la ubicación de preferencia (5 horas)

1.2 Generación de la base de datos de videos (30 horas)

1.2.1 Grabación de videos con correspondencia "alarma tipo 1" (10 horas)

1.2.2 Grabación de videos con correspondencia "alarma tipo 2" (10 horas)

1.2.3 Grabación de videos con correspondencia "alarma tipo 3" (10 horas)

1.3 Documentación (20 horas)

1.3.1 Documentar detalles de la base de datos - nombre del archivo, duración, tipo de alarma, etc. (5 horas)

1.3.2 Redactar informe de Incremento 1 (15 horas)

2. Incremento 2 (200 horas)

2.1 Estudio sobre conceptos relacionados con características de Haar y la biblioteca Dlib. (30 horas)

2.2 Instalación de la biblioteca y análisis de las funciones pertinentes de la misma (20hs)

2.3 Desarrollo de sistema de detección de pérdidas de atención basándonos en la base de datos generada previamente. (100 horas)

2.4 Pruebas exhaustivas (30 horas)

2.5 Documentar (20 horas)

2.5.1 Documento con detalles sobre el software (5 horas)

2.5.2 Informe de Incremento 2 (15 horas)

3. Incremento 3 (100 horas)

3.1 Desarrollo de método para la utilización de la Dinámica temporal de los videos. (30 horas)

3.2 Asignar en el software alarmas correspondientes a las detecciones pertinentes(50 horas)

3.3 Documentar (20 horas)

3.3.1 Documento con detalles sobre las alarmas y porcentaje de acierto en los videos. (5 horas)

3.3.2 Informe de Incremento 3 (15 horas)

4. Incremento 4 (140 horas)

4.1 Estudiar conceptos de desarrollo para interfaces (20 horas)

4.2 Desarrollo de interfaz para una comunicación amigable con el cliente. (50 horas)

4.3 Vinculación con el software desarrollado en el incremento anterior.(50 horas)

4.4 Documentar (20 horas)

4.4.1 Redacción de manual de usuario (5 horas)

4.4.2 Informe de incremento 4 (15 horas)

Total horas: 500 hs/hombre

Cronograma

Para la realización del mismo, se tomó como fecha de inicio del proyecto el 15/1/2018. Ésta fecha tiene su justificación en un intercambio a Chile llevado a cabo por el alumno responsable del desarrollo del proyecto (Fecha inicio: 9/08/2017 Hasta: 23/12/2017). Así, sabiendo la fecha de inicio, y sabiendo que tendremos que cumplir 100 períodos de trabajo suponiendo que se trabajará 5 horas al día 5 días a la semana. Se estima que la fecha de finalización del proyecto sería el 27/6/2018. La imagen del cronograma se adjunta en la página 9.

Puntos de control a presentar a la cátedra

Los puntos de control considerados para este proyecto, estarán vinculados con los incrementos del mismo y servirán para ir midiendo el avance del proyecto por parte de la cátedra. Se detallan a continuación detalles y fechas de los mismo.

Informe de avance nº 1: Informe de detalles y conclusiones del incremento 1.

Fecha de entrega: 15/02/2018

Descripción: Informe en los que quedará explícito el contenido de la base de datos de videos con la que se trabajará, junto con características de utilidad de cada uno, duración, tipo de alarma asignada, etc. Así como también conclusiones pertinentes a las actividades realizadas en el incremento 1.

Informe de avance nº 2: Informe de detalles y conclusiones del incremento 2.

Fecha de entrega: 27/03/2018

Descripción: Se planea la entrega de un documento con las especificaciones del software desarrollado en el cual se describan las funcionalidades logradas y queden adjuntas todas conclusiones obtenidas a partir de las actividades ejercidas en el incremento 2.

Informe de avance nº 3: Informe de detalles y conclusiones del incremento 3.

Fecha de entrega: 18/05/2018

Descripción: Basándonos en el software entregado anteriormente, se entrega la documentación de la actualización realizada donde estarán especificados los detalles de las distintas alarmas desarrolladas y las conclusiones obtenidas en el incremento 3.

Informe de avance nº 4: Informe de detalles y conclusiones del incremento 4.

Fecha de entrega: 27/06/2018

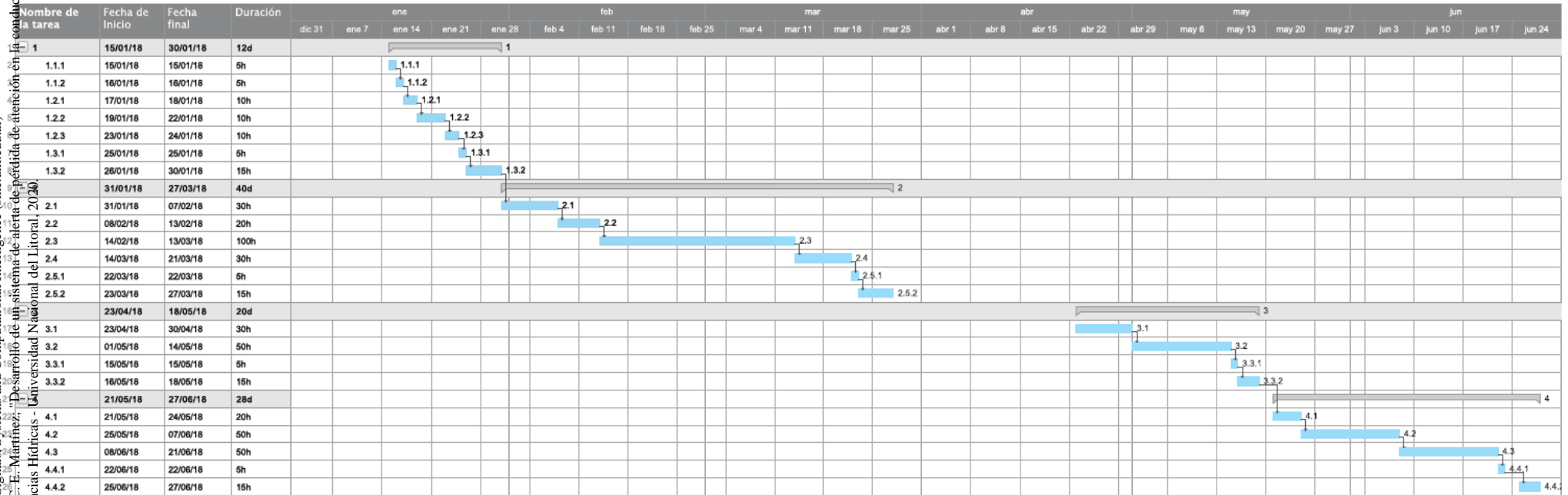
Descripción: Se contemplará un documento detallando la interfaz realizada para el uso del cliente y/o desarrollador junto con la entrega de un manual de usuario con los detalles puntuales sobre su funcionalidad. También se entregarán las conclusiones del incremento mencionado, así como también conclusiones generales del proyecto finalizado.

Criterios de aceptación

Además de el cumplimiento de los objetivos planteados anteriormente, su aceptación va a estar ligada a la corrección del informe final por parte de los directores del proyecto.

Diagrama de Gantt

Proyecto Final de Carrera: Cronograma



Link al cronograma: https://drive.google.com/open?id=0B-h7XL4P7Js_eEtvNIJEUTRVMk0

Riesgos

En esta sección se presentarán los riesgos identificados que pueden afectar al proyecto en sí. A partir de la prioridad e impacto de los mismos se definirá la estrategia a seguir, tanto en lo que respecta a la respuesta a los mismos como a su contingencia en caso de que ocurriesen.

- **Equipo de trabajo:**

Falta de disponibilidad del director

Descripción: Debido a situaciones personales o laborales, que el director no esté disponible en algún momento del proyecto.

Probabilidad: Baja.

Impacto: Medio.

Mitigación: Realizar varias consultas que demanden poco tiempo, en lugar de pocas que demanden mucho tiempo.

Contingencia: Considerar al co-director para que desempeñe el cargo del director.

Falta de disponibilidad del responsable de la ejecución del proyecto

Descripción: Demora en la ejecución del proyecto por inconvenientes de carácter personal.

Probabilidad: Baja.

Impacto: Medio.

Mitigación: Definir márgenes de tiempo razonable para cada etapa del proyecto destinados a la contingencia de dichos inconvenientes.

Contingencia: Reajustar las tareas del proyecto a partir de una redefinición del cronograma acorde a los tiempos que el responsable se ausentó.

- **Recursos:**

Insumos defectuosos

Descripción: Que los insumos adquiridos tengan algún defecto de fabricación.

Probabilidad: Baja.

Impacto: Alto.

Mitigación: Realizar consultas con expertos en el tema acerca de qué proveedores son más confiables y reconocidos.

Contingencia: Utilizar la garantía de los insumos para obtener nuevos sin defectos.

Falta de disponibilidad de recursos esenciales para la grabación

Descripción: Debido a razones que no se pueden controlar, que al momento de grabar falten elementos esenciales para la grabación de los videos, se habla por supuesto de cámara, el automóvil, y hasta los conductores mismos.

Probabilidad: Baja.

Impacto: Alto.

Mitigación: Planear y revisar los equipos de forma previa con el fin de disminuir la probabilidad de que el riesgo ocurra.

Contingencia: Ante la ausencia de los recursos, reorganizar una futura sesión de grabación fuera del tiempo planificado con el fin de reponer el tiempo perdido

- **Información/Conocimiento:**

Complicaciones con el manejo de nuevas tecnologías

Descripción: Retraso en las actividades de diseño y desarrollo debido a inconvenientes con el manejo de las herramientas utilizadas para dicho fin.

Probabilidad: Baja.

Impacto: Medio.

Mitigación: Capacitación mediante bibliografía autorizada y cursos online acerca de las herramientas de desarrollo a utilizar.

Contingencia: Consultar a los directores sobre el tema y sobre qué fuentes confiables recomiendan para adquirir los conocimientos en dichas áreas.

Costos

COSTO	CANTIDAD	PRECIO U	SUBTOTAL
Bienes de capital			
Notebook*	1	950	5700
Cámara GoPro*	1	600	3600
Soporte físico para la cámara.*	1	100	600
Herramientas varias			800
Consultas			
Consultas sobre utilización de cámara GoPro.	4Hs	100	400
Consultas a miembros del Sinc(i)	15Hs	300	4500
RRHH			
Horas Programador	500Hs	130	65000
Horas Director	50Hs	500	25000
Horas Co-Director	50Hs	500	25000
Horas conductores	30Hs	550	16500
Otros			
Acceso a internet	6 meses	300	1800
Energía eléctrica	$(0,525\text{kw}/\text{Hs}) * 300\text{Hs} = 157,5\text{kw}$	0,91	143,80
Impresión y anillado de informe final	3	300	900
Impresión entregables	4	50	200
Nafta	41 lts.	22	902
		TOTAL	\$ 151.045,80

* Se calcularon amortizaciones mensuales por los meses de utilización del bien de capital. Tomando 6 meses como el tiempo de trabajo.

Recursos

- Recursos existentes:

Hardware:

- Notebook Lenovo Thinkpad T430 Intel Core i5, 8gb RAM, 500 gb Disco Rigido, 14" pantalla.
- Automóvil Nissan March. Modelo 2016.
- Cámara GoPro Hero.
- Soporte para vehículos para la cámara.

Software:

- Entorno de desarrollo: Zinjai.
- Sistema Operativo: Linux y Windows 10.

-Entorno para redacción de informes: Microsoft Word 2007.

Recursos Humanos:

- Alumno que realiza el proyecto.
- Director y co-director del proyecto.
- Conductores para la grabación de videos

Insumos y servicios:

- Insumos de oficina: lapicera, hojas, etc.
- Infraestructura necesaria para desarrollar el proyecto.
- Servicio de internet.
- Servicio de electricidad.
- Nafta para suministrar el automóvil.

Bibliografía

Project management institute, inc (2013): "Guía de los Fundamentos para la Dirección de Proyectos". PMI Book. Cuarta edición

Ander Egg, E, Aguilar Idañez, M. (1998): "Cómo elaborar un proyecto". Editorial Lumen.

Verzuh E. (2015): "The Fast Forward MBA in Project Management, 5th Edition". Editorial Wiley. Quinta edición

[1] Luchemos por la vida. Estadísticas: Muertes en Argentina durante 2016. URL
Accedida el 13 de abril de 2017:

<http://www.luchemos.org.ar/es/estadisticas/muertosanuales/muertos-en-argentina-durante-2016>

[2] Diario Clarín: "Argentina, a la cabeza de un ranking de muertes en el tránsito". URL
accedida el 13 de abril de 2017:

https://www.clarin.com/sociedad/itf-ocde-siniestros_viales-estadistica_internacional_0_ryfsoOFv7g.html

[3] Centro de Experimentación, Investigación y Seguridad Vial (CESVI) - Estadísticas sobre Seguridad Vial: Accidentes relevados por causa, accidentes relevados por tipo de vehículo, fallas humanas más comunes. URL accedida el 24 de abril de 2017:

<https://www.cesvi.com.ar/SeguridadVial/Estadisticas/SeguridadEstadisticas.aspx>

[4] Ministerio del Interior. Subdirección general de intervención y políticas viales. Unidad de intervención educativa. "Distracciones al volante". URL accedida el 13 de abril de 2017:

http://www.dgt.es/PEVI/documentos/catalogo_recursos/didacticos/did_adultas/Distracciones_al_volante.pdf

[5] Actualidad Motor: "Bosch desarrolló Sistema de Detección del Sueño". URL accedida el 13 de abril de 2017:

<https://www.actualidadmotor.com/bosch-desarrollo-sistema-de-deteccion-del-sueno/>

[6] Universidad de Barcelona. "Curso de Detección de Objetos: Cascada de Clasificadores"

URL accedida el 13 de abril de 2017:

<https://www.coursera.org/learn/deteccion-objetos/lecture/pRnHu/l5-5-cascada-de-clasificadores>

[7] Pressman, R. S., & Troya, J. M. (1988). *Ingeniería del software* (No. 001.64 P74s.). McGraw Hill.

[8] Librería Dlib. Ingresado el 22/05/2017. URL : <http://Dlib.net/>

[9] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer-Verlag New York

[10] Laganier, R. (2017) . *OpenCV 3 Computer Vision Application Programming Cookbook, 3rd Edition*. Packt Publishing

[11] Facial landmarks using Dlib, opencv, and python. Ingresado el 22/05/2017. URL accedida el 3 de Mayo de 2017: <http://www.pyimagesearch.com/2017/04/03/facial-landmarks-Dlib-opencv-python/>



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas

PROYECTO FINAL DE CARRERA
INGENIERÍA INFORMÁTICA

Manual de Usuario

**Desarrollo de un sistema de detección de pérdidas de atención en la
conducción basado en visión computacional**

Alumno: Farré, Gonzalo

Director: Martínez, César

Co-Director: Albornoz, Enrique Marcelo

Índice general

1. Introducción al sistema.....	3
1.1 Instalaciones generales y requisitos básicos	3
1.2 Instalación del IDE Atom	6
2. Archivos del proyecto	7
2.1 Configuración de los archivos	7
3. Compilación y ejecución del sistema	11
3.1 Compilación	11
3.2 Ejecución.....	12
Referencias	16

1. Introducción al sistema

Se presenta a continuación un manual de usuario básico con el fin de que desarrolladores y estudiantes del área en cuestión puedan investigar el funcionamiento del método, y analizar sus propios videos proponiendo cambios y optimizaciones.

1.1 Instalaciones generales y requisitos básicos

Para lograr el desarrollo del sistema se utilizó una Notebook Lenovo Thinkpad T430 Intel Core i5, 8gb RAM, 500 gb Disco Rígido. Por lo que se recomienda un equipo de características similares para la ejecución y desarrollo del sistema.

Además de los requisitos físicos, es necesario la instalación de Ubuntu 16.04.5, la biblioteca Open CV 3.4.3, el compilador CMake 3.13, la biblioteca Dlib 19.16 y el editor de texto Atom version 1.29 (Ver Figura 1.1). Cada uno de éstos componentes constituyen lo que viene a ser el soporte de sistema y los cuales debemos tener instalados en nuestra PC para la correcta compilación y ejecución del mismo. Se detallan los links de descarga en las referencias. [1]

Al ser desarrolladores los usuarios principales de utilización del sistema, se obviarán detalles sobre instalaciones ya que las mismas son totalmente standard y no requieren conocimiento específico de ningún tipo.

Release Candidate (3.13.0-rc3)

The release was packaged with CPack which is included as part of the release. The .sh files are self extracting gzipped tar files. To install a .sh file, run it with /bin/sh and follow the directions. The OS-machine.tar.gz files are gzipped tar files of the install tree. The OS-machine.tar.Z files are compressed tar files of the install tree. The tar file distributions can be untared in any directory. They are prefixed by the version of CMake. For example, the Linux-x86_64 tar file is all under the directory cmake-Linux-x86_64. This prefix can be removed as long as the share, bin, man and doc directories are moved relative to each other. To build the source distributions, unpack them with zip or tar and follow the instructions in Readme.txt at the top of the source tree. See also the [CMake 3.13 Release Notes](#). Source distributions:

Platform	Files
Unix/Linux Source (has \n line feeds)	cmake-3.13.0-rc3.tar.gz cmake-3.13.0-rc3.tar.Z
Windows Source (has \r\n line feeds)	cmake-3.13.0-rc3.zip

Binary distributions:

Platform	Files
Windows win64-x64 Installer: Installer tool has changed. Uninstall CMake 3.4 or lower first!	cmake-3.13.0-rc3-win64-x64.msi
Windows win64-x64 ZIP	cmake-3.13.0-rc3-win64-x64.zip
Windows win32-x86 Installer: Installer tool has changed. Uninstall CMake 3.4 or lower first!	cmake-3.13.0-rc3-win32-x86.msi
Windows win32-x86 ZIP	cmake-3.13.0-rc3-win32-x86.zip

(I)

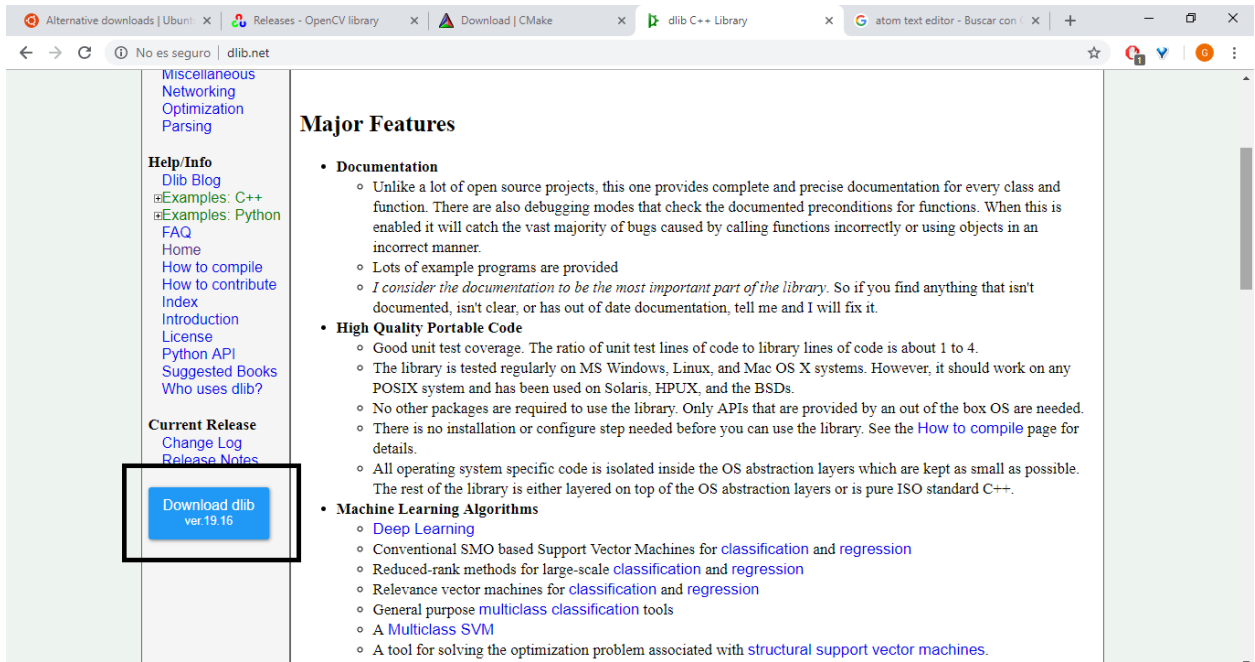
BitTorrent

BitTorrent is a peer-to-peer download network that sometimes enables higher download speeds and more reliable downloads of large files. You will need to install a BitTorrent client on your computer in order to enable this download method.

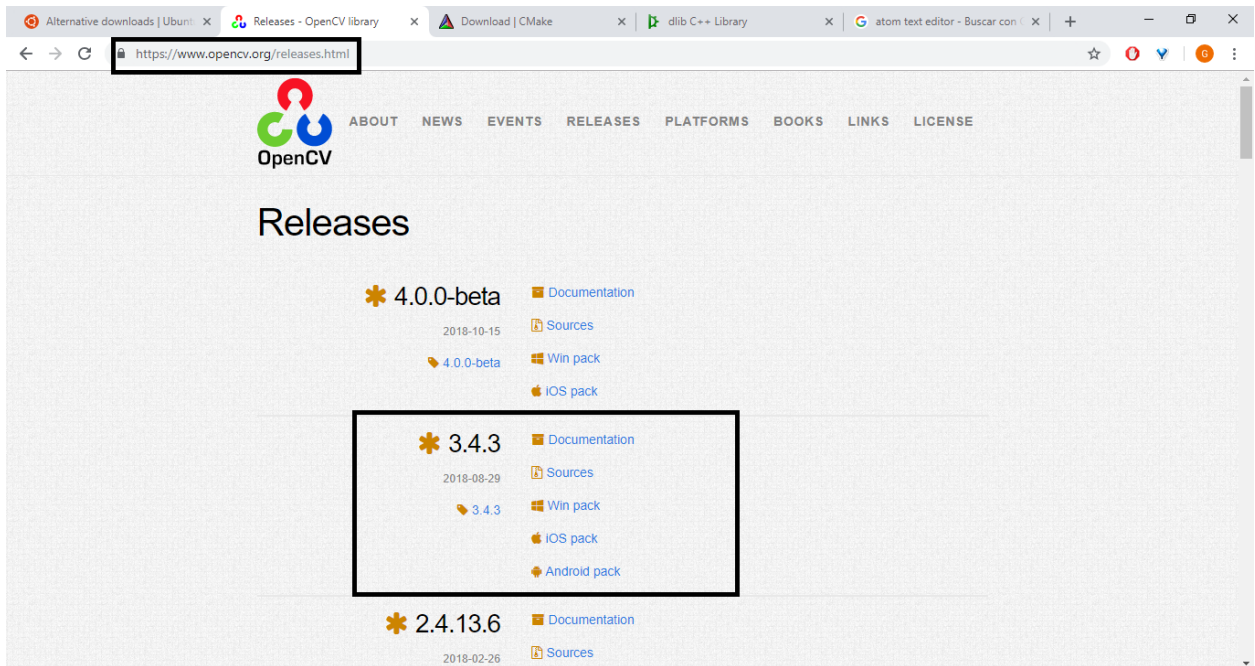
Ubuntu 18.10	Ubuntu 18.04.1 LTS	Ubuntu 16.04.5 LTS	Ubuntu 14.04.5 LTS
Ubuntu 18.10 Desktop (64-bit)	Ubuntu 18.04.1 Desktop (64-bit)	Ubuntu 16.04.5 Desktop (64-bit)	Ubuntu 14.04.5 Desktop (64-bit)
Ubuntu 18.10 Server (64-bit)	Ubuntu 18.04.1 Server (64-bit)	Ubuntu 16.04.5 Desktop (32-bit)	Ubuntu 14.04.5 Desktop (32-bit)
		Ubuntu 16.04.5 Server (64-bit)	Ubuntu 14.04.5 Server (64-bit)
		Ubuntu 16.04.5 Server (32-bit)	Ubuntu 14.04.5 Server (32-bit)

Alternative Ubuntu | Other images and | Past releases and

(II)



(III)

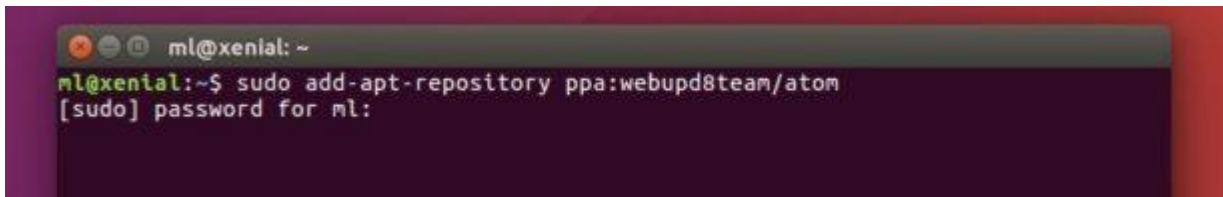


(IV)

Fig. 1.1: I, II, III, IV Entornos pertenecientes a las distintas descargas requeridas por el sistema.

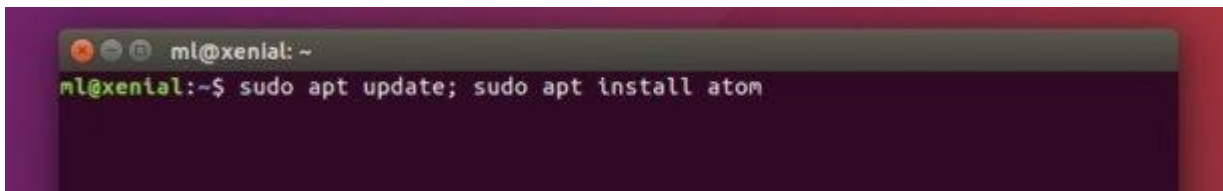
1.2 Instalación del IDE Atom

Al no presentarse una descarga standard para el ambiente de desarrollo, se deberá instalar el IDE mediante consola. Cabe mencionar que se debe ejecutar con el comando sudo para lograr una ejecución del tipo usuario Root, y se facilita además el comando necesario para actualizar el Atom a la última versión en caso de ya tenerlo instalado. (Ver Figura 1.2)



```
ml@xenial: ~  
ml@xenial:~$ sudo add-apt-repository ppa:webupd8team/atom  
[sudo] password for ml:
```

(I)



```
ml@xenial: ~  
ml@xenial:~$ sudo apt update; sudo apt install atom
```

(II)

Fig. 1.2: I, II Comandos pertenecientes tanto a la instalación como actualización de Atom en el sistema operativo Ubuntu.

2. Archivos del proyecto




Una vez instalado todos los programas requeridos por el sistema, se procede a descargar los archivos que corresponden al sistema en sí. Se brinda en la sección referencias, el link correspondiente a la carpeta de Google Drive [2]

2.1 Configuración de los archivos

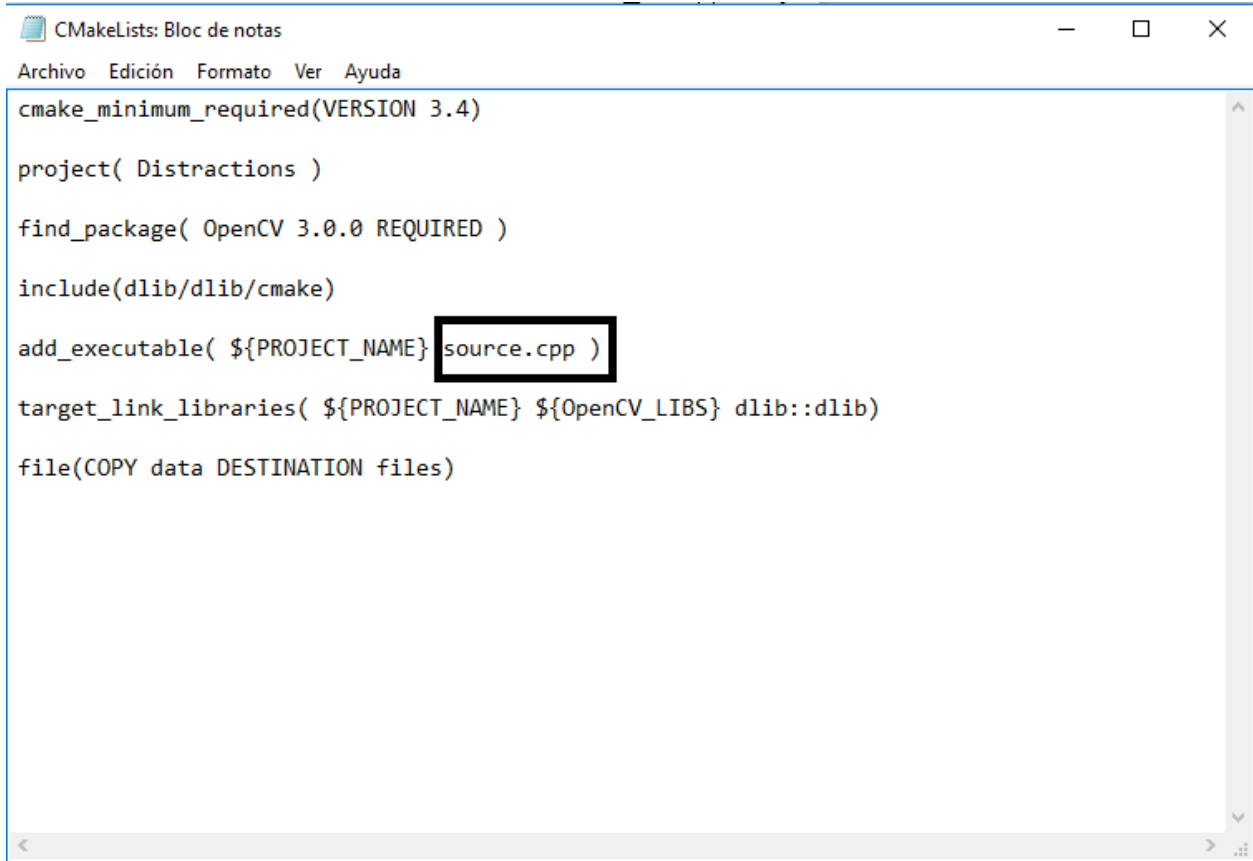
Una vez descargados los archivos, se debe verificar que el archivo de compilación, llamado CMakeLists.txt, haga referencia directamente al archivo que se busca ejecutar. En nuestro caso, Source.cpp. (Ver figura 2.1)

Luego de dicha verificación, se procede a abrir el archivo Source.cpp con la intención de verificar que todas las librerías estén cargadas de forma adecuada.

Mi unidad > Distractions ▾

Nombre ↑	Propietario	Última modificación	Tamaño
 data	yo	19:17 yo	—
 CMakeLists.txt	yo	19:16 yo	297 bytes
 source.cpp	yo	19:16 yo	38 KB

(I)



```
cmake_minimum_required(VERSION 3.4)

project( Distractions )

find_package( OpenCV 3.0.0 REQUIRED )

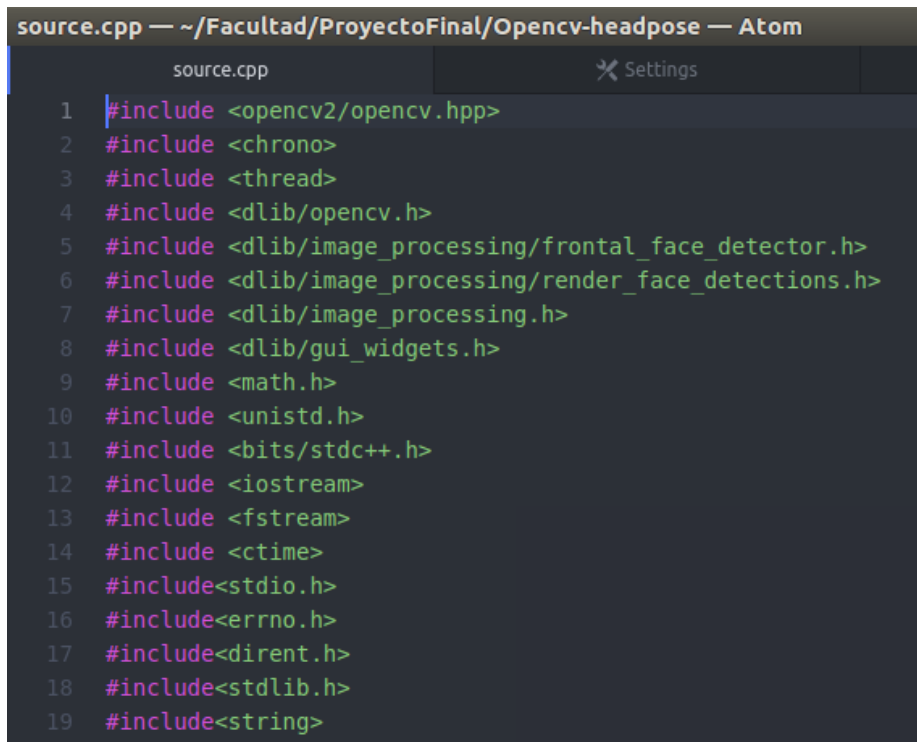
include(dlib/dlib/cmake)

add_executable( ${PROJECT_NAME} source.cpp )

target_link_libraries( ${PROJECT_NAME} ${OpenCV_LIBS} dlib::dlib)

file(COPY data DESTINATION files)
```

(II)



```
source.cpp — ~/Facultad/ProyectoFinal/Opencv-headpose — Atom
source.cpp Settings
1 #include <opencv2/opencv.hpp>
2 #include <chrono>
3 #include <thread>
4 #include <dlib/opencv.h>
5 #include <dlib/image_processing/frontal_face_detector.h>
6 #include <dlib/image_processing/render_face_detections.h>
7 #include <dlib/image_processing.h>
8 #include <dlib/gui_widgets.h>
9 #include <math.h>
10 #include <unistd.h>
11 #include <bits/stdc++.h>
12 #include <iostream>
13 #include <fstream>
14 #include <ctime>
15 #include <stdio.h>
16 #include <errno.h>
17 #include <dirent.h>
18 #include <stdlib.h>
19 #include <string>
```

(III)

Fig 2.1: (I) Carpeta de Google Drive con los archivos de a descargar (II) Configuración correcta del archivo CMakeLists.txt (III) Configuración correcta de librerías a utilizar.

Lo que hace falta mencionar ahora, será la colocación de videos para el posterior análisis en el sistema. Para esto se deben colocar en la carpeta central del sistema – donde se encuentra source.cpp – con los nombres clave que se han descrito en el informe presentado anteriormente. (Ver Figura 2.2)

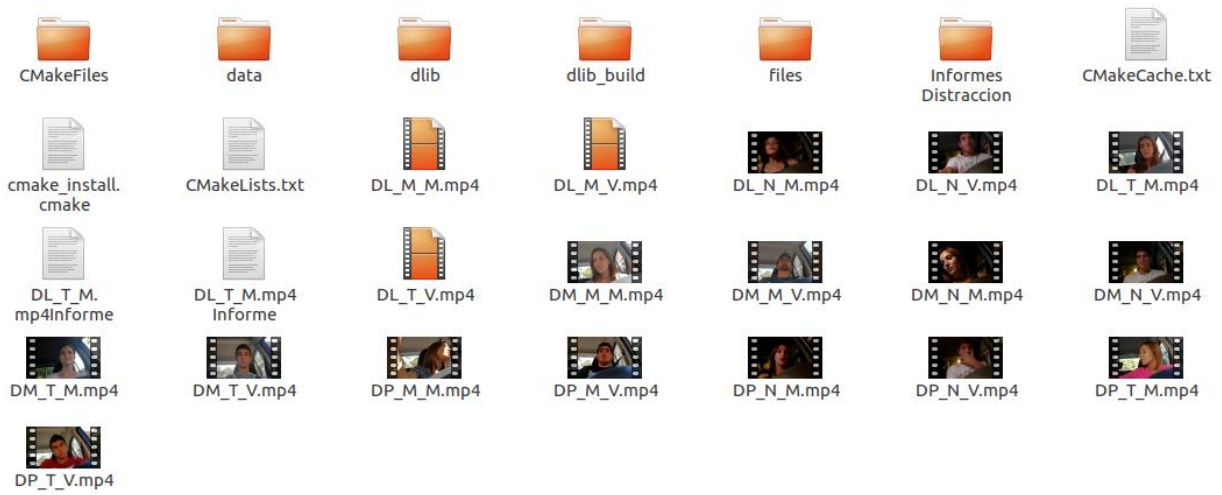


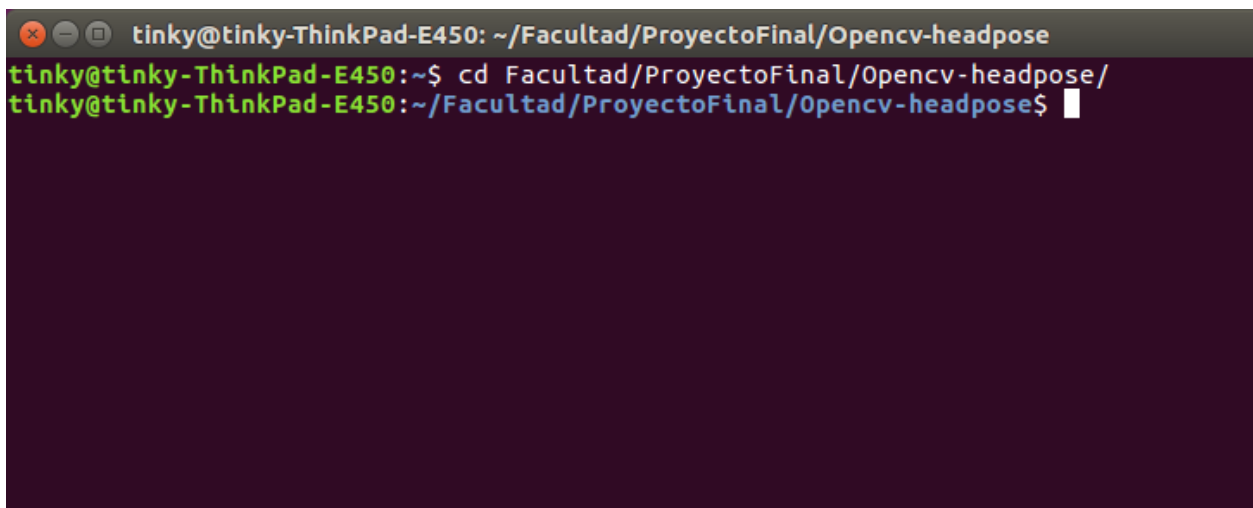
Fig. 2.2: Carpeta central del sistema con los videos a evaluar

3. Compilación y ejecución del sistema

Para la compilación y ejecución finalmente del sistema, se deberá recurrir a CMake de forma que todos los cambios realizados en el archivo Source.cpp sean cargados, y estén listos para la ejecución.

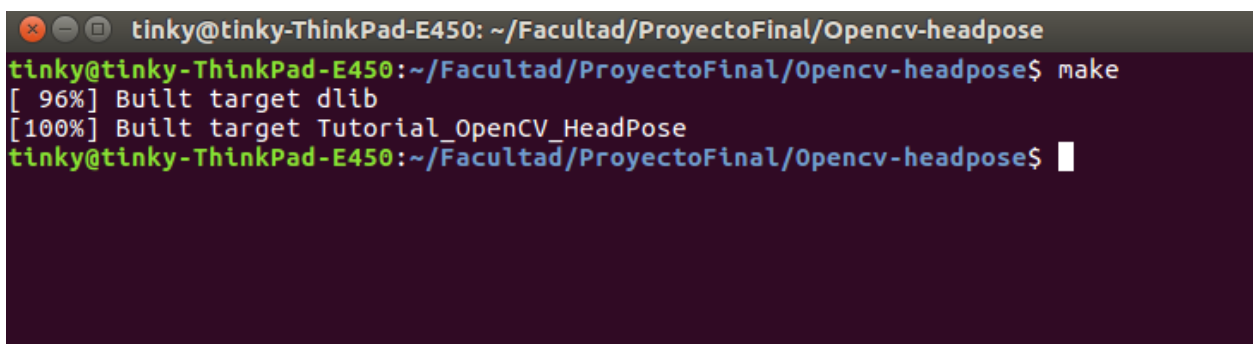
3.1 Compilación

Para utilizar CMake, basta con trazar la ruta de la terminal hasta donde se encuentra el archivo Source.cpp, y luego ejecutar el comando Make. De esta manera, y si todo está en orden, estaremos compilando el archivo y dejándolo listo para la ejecución. (Ver Figura 3.1)



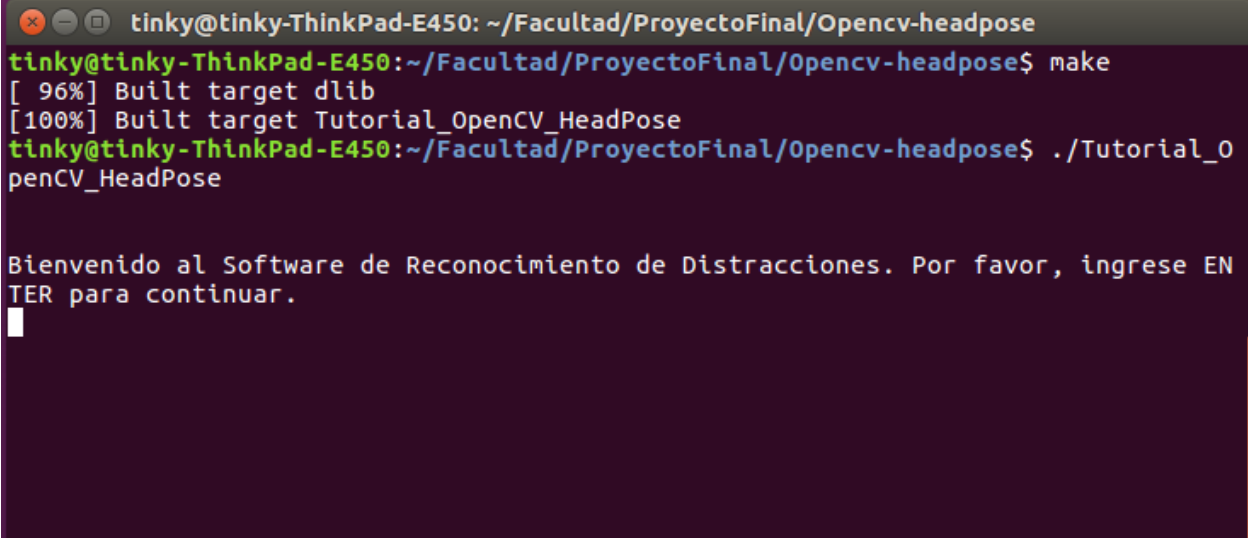
```
tinky@tinky-ThinkPad-E450: ~/Facultad/ProyectoFinal/Opencv-headpose
tinky@tinky-ThinkPad-E450:~$ cd Facultad/ProyectoFinal/Opencv-headpose/
tinky@tinky-ThinkPad-E450:~/Facultad/ProyectoFinal/Opencv-headpose$
```

(I)



```
tinky@tinky-ThinkPad-E450: ~/Facultad/ProyectoFinal/Opencv-headpose
tinky@tinky-ThinkPad-E450:~/Facultad/ProyectoFinal/Opencv-headpose$ make
[ 96%] Built target dlib
[100%] Built target Tutorial_OpenCV_HeadPose
tinky@tinky-ThinkPad-E450:~/Facultad/ProyectoFinal/Opencv-headpose$
```

(II)



```
tinky@tinky-ThinkPad-E450: ~/Facultad/ProyectoFinal/Opencv-headpose
tinky@tinky-ThinkPad-E450:~/Facultad/ProyectoFinal/Opencv-headpose$ make
[ 96%] Built target dlib
[100%] Built target Tutorial_OpenCV_HeadPose
tinky@tinky-ThinkPad-E450:~/Facultad/ProyectoFinal/Opencv-headpose$ ./Tutorial_0
penCV_HeadPose

Bienvenido al Software de Reconocimiento de Distracciones. Por favor, ingrese EN
TER para continuar.
█
```

(III)

Fig. 3.1: (I) Redireccionamiento de la consola a la carpeta raíz. (II) Compilación del archivo mediante comando Make (III) Ejecución del archivo correspondiente.

3.2 Ejecución

Con respecto al uso del sistema, el mismo fue diseñado para ser intuitivo y no dejar lugar a la ambigüedad. Con una pantalla de bienvenida inicia ordenando el ingreso de una tecla para continuar con la ejecución, como se mostró en la imagen anterior.

Luego de haber confirmado mediante el ingreso de la tecla, se da a elegir el video a analizar, poniendo en pantalla los 18 videos predefinidos. Cabe mencionar que se dan a elegir 19 opciones, siendo esta ultima la opción de ingresar un video extra. Éste video deberá estar en la carpeta central del sistema, y deberá ser buscado con dicho nombre a la hora de la ejecución. (Ver Figura 3.2)

```

A continuación, ingrese el video a evaluar en base a los siguientes items:
1) Distracción leve / Mañana / Varón
2) Distracción leve / Mañana / Mujer
3) Distracción leve / Tarde / Varón
4) Distracción leve / Tarde / Mujer
5) Distracción leve / Noche / Varón
6) Distracción leve / Noche / Mujer

7) Distracción Moderada / Mañana / Varón
9) Distracción Moderada / Tarde / Varón
10) Distracción Moderada / Tarde / Mujer
11) Distracción Moderada / Noche / Varón
12) Distracción Moderada / Noche / Mujer

8) Distracción Moderada / Mañana / Mujer
13) Distracción Noche / Mañana / Varón
14) Distracción Noche / Mañana / Mujer
15) Distracción Noche / Tarde / Varón
16) Distracción Noche / Tarde / Mujer
17) Distracción Noche / Noche / Varón
18) Distracción Noche / Noche / Mujer

19) Reproduzca un archivo a elección

```

Fig. 3.2: Listado de videos sugeridos por el sistema

Una vez escogido el video a través de su número, el sistema dará inicio al análisis del video, detectando las distracciones que se presentan en el mismo. (Ver Figura 3.3) [3]

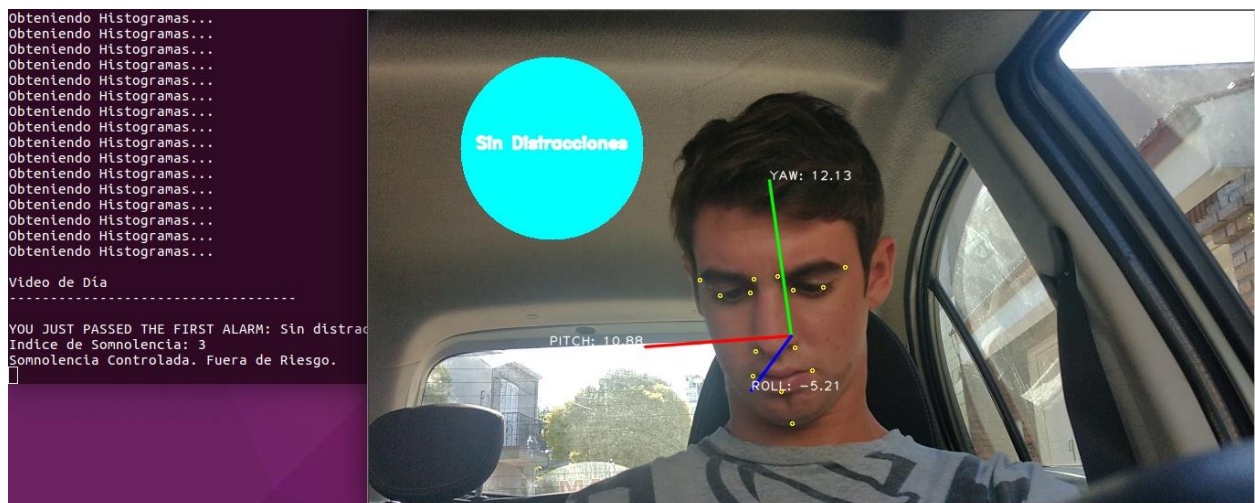
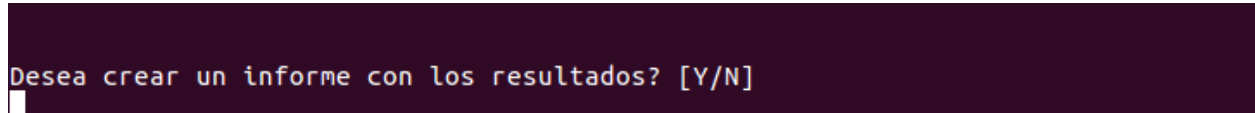


Fig. 3.3: Sistema en funcionamiento con video de tarde, varon.

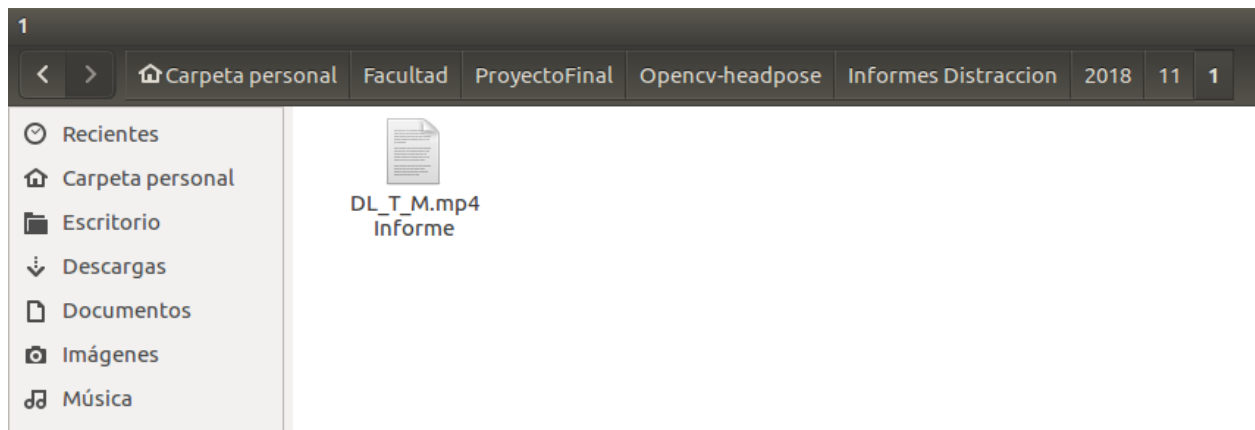
Luego de finalizado el programa, el sistema brinda la opción de elegir si se desean guardar los resultados en un archivo, o si se desean ver los resultados por consola. Si se elige la opción del archivo, el sistema creará una serie de directorios

correspondientes a la fecha del día, y guardará el archivo con el nombre: <Nombre del Video> Informe. (Ver Figura 3.4)



Desea crear un informe con los resultados? [Y/N]

(I)



(II)


```

*DL_T_M.mp4 Informe (~:/Facultad/ProyectoFinal/Opencv-headpose/Informes Distraccion/2018/10/17) - gedit
Abrir ▾ [?]
INFORME DE DISTRACCIONES|
Video DL_T_M.mp4

Frames en total: 1854
Primera alarma en segundo : 16
Segunda alarma en segundo: 33
Tercera alarma en segundo: 49
-----
Frames no detectados: 14
Duración Video: 66 segundos
Índice de Distracciones: 71
Total de Distracciones: 7
-----
Detalle de Distracciones:
-----
Distracción 1: Distraccion Parcial en segundo: 3
Distracción 2: Distraccion Parcial en segundo: 4
Distracción 3: Distraccion Parcial en segundo: 14

Hasta ahora se trató de distracciones LEVES

Distracción 4: Distraccion Parcial en segundo: 20
Distracción 5: Distraccion Parcial en segundo: 21

Hasta ahora se trató de distracciones LEVES

Distracción 6: Distraccion Parcial en segundo: 46

Hasta ahora se trató de distracciones LEVES

Distracción 7: Distraccion Parcial en segundo: 53

Hasta ahora se trató de distracciones LEVES
-----
Video Finalizado. Gracias por reproducir.

```

(III)

Fig. 3.4: (I) Diálogo para la creación de un informe de Distracciones. (II) Nombre y ubicación del archivo de distracciones (III) Archivo de Distracciones.

Referencias

[1] Link de descarga de OpenCV 3.4.3 Accedida el 16/11/18:
<https://www.opencv.org/releases.html>

Link de descarga de Ubuntu 16.04.5 Accedida el 16/11/18:
<https://www.ubuntu.com/download/desktop>

Link de descarga de CMake 3.13 Accedida el 16/11/18:
<https://cmake.org/download/>

Link de descarga de Dlib 19.16 Accedida el 16/11/18:
<http://Dlib.net/files/Dlib-19.16.zip>

[2] Link de carpeta de Google Drive con archivos necesarios, Accedida el 16/11/18:
<https://drive.google.com/open?id=1wNP0qSUzloQus7-zrxfLw1-fvJnq1F1i>

[3] Link directo a videos de ejemplo junto su funcionamiento, Accedida el 16/11/18:
https://drive.google.com/open?id=105Wt_bxBa6YhwFp8DDIYUNZCd6vkID1R