

Generación de Servicios Digitales en Ciudades Inteligentes a Partir de las Capacidades de los Sistemas de Cámaras

J. P. D'Amato^{1,2}, L. Dominguez^{1,2}, A. Perez^{1,3}, A. Rubiales^{1,3}, F. Stramana^{1,3}

Juan.damato@gmail.com

¹ Instituto PLADEMA – UNCPBA, Campus Argentino, 7000, Tandil, Argentina.

² Consejo Nacional de Investigaciones Científicas y Técnicas, Av. Rivadavia 1917, CABA, Argentina.

³ Comisión de Investigaciones Científicas, Calle 526 e/10 y 11, La Plata, Argentina.

Pages: 566–578

Resumen: En la mayoría de las ciudades del mundo, existen centros de monitoreo con video cámaras solo utilizadas para la vigilancia, buscando ilícitos o accidentes. Pero esta inmensa cantidad de información que se obtiene a partir de videos (cientos de GBs por días), es posible aprovecharla con otros fines. Los gobiernos de dichas ciudades están buscando soluciones para ofrecer diferentes servicios, tales como, estacionamiento inteligente, manejo de semáforos, congestionamientos, entre otros. Dada la infraestructura ya disponible de cámaras, hoy se puede pensar en que estos servicios sean basados en tecnología imágenes en pos de una ciudad inteligente. En este trabajo se evalúan la potencialidad del uso de datos de imágenes para generar nuevos indicadores que definan servicios para los ciudadanos, en base a una plataforma unificada y abierta. Se realiza una propuesta de consideraciones técnicas para que esto sea factible, haciendo un estudio del esfuerzo computacional necesario y una caracterización de los tipos de servicios que se pueden crear, así como una propuesta arquitectónica.

Palabras-clave: ciudad inteligente; video-seguridad; visión computadora.

Generation of digital services for smart-cities exploiting surveillance systems capabilities

Abstract: In most cities in the world, there are monitoring centers with video cameras only used for video surveillance, looking for illicit or accidents. But this immense amount of information that comes from videos (hundreds of GBs per day), could be used for other purposes. At the same time, the governments of these cities are looking for new different services, such as smart parking, traffic light management, congestion, among others. Given the already available camera infrastructure, new services based on image technology arise. This paper evaluates the potential of using image data to generate indicators necessary to define services for citizens, based on an unified and open platform. A proposal of technical

considerations is made to show the feasibility. Also, a study of the necessary computational effort and a characterization of services types that can be created is presented.

Keywords: smartCities: surveillance systems; computer vision.

1. Introducción

Hoy, la seguridad sigue siendo uno de los temas principales en las políticas de gobierno de todo el mundo. La preocupación que generan en la sociedad los altos índices de violencia e inseguridad actuales (Kessler, 2009) ha motivado el estudio de nuevas tecnologías que permitan a los habitantes vivir con mayor tranquilidad. Este hecho se intensifica en grandes ciudades, donde se producen situaciones lamentables con mayor frecuencia. Según Dammert et al. (2010), entre el año 2000 y 2008, la tasa de población penitenciaria en América Latina creció un 42. Por otro lado, estos sistemas han generado un efecto adverso: las personas se sienten siempre “observadas”, con un impacto negativo ya que ven afectados sus derechos de privacidad.

Los sistemas de seguridad basadas en videos son una realidad y una característica determinante en las denominadas “Smart cities” o “ciudades inteligentes”. Cuando hablamos de una Smart City nos referimos a un tipo de desarrollo urbano basado en la sostenibilidad, por lo tanto, se trata de una calificación que se hace a las ciudades en medida de sus inversiones en aspectos sociales, infraestructuras de energía y transporte, tecnologías de la comunicación y todas aquellas acciones que contemplen y promuevan una calidad de vida elevada, basada en un desarrollo económico-ambiental sostenible. En la definición de ciudades inteligentes (Smart City Framework), se plantean cinco (5) ejes de trabajo, dentro de la que se encuentra “video surveillance”, tal como se muestra en la Figura 1.



Figura 1 – Diferentes ejes y servicios de una plataforma de “ciudad inteligente”

Esta transformación de las ciudades hacia el futuro se ve especialmente afectada por la revolución tecnológica que llevamos viviendo. Se estima que, en el próximo año, habrá 50.000 millones de dispositivos conectados a Internet. Para que esto sea posible, las ciudades tendrán que invertir en esta digitalización para poder adaptarse a las nuevas necesidades de una sociedad cada vez más conectada y tendrán que lidiar con la cantidad de información generada. A su vez, se estima que la demanda del mercado mundial de Ciudades Inteligentes que podría alcanzar un valor de 1,5 trillones de dólares. En esta misma línea, las Ciudades Inteligentes, en base al Internet de las Cosas, contarán con 10.000 millones de elementos conectados en el año 2020, es decir, casi 10 veces más que los elementos conectados previstos para este año 2015, según la predicción de Gartner. (ONTSI 2015).

Esta nueva realidad tiene diferentes desafíos tecnológicos que deben ser atacados a fin de dar respuesta a una sociedad basada en la información: manipulación de gran cantidad de datos y disponibilidad de los mismos. Para dar una idea, en una cámara localizada en un lugar con tránsito vehicular/peatonal moderado, sólo en el lapso de una hora, es posible identificar aproximadamente 500 objetos. Si a esto lo escalamos por la cantidad de cámaras disponibles en una ciudad, en un mes identificamos más de 6×10^7 objetos. Esta información puede ser utilizada para el control óptimo de semáforos (D'Amato et al 2011).

Entendiendo que los sistemas de video-vigilancia son una pieza clave en la estructura de una ciudad inteligente, no solo en el ámbito de la seguridad sino también en el social, es que este grupo ha trabajado y generado diferentes trabajos en esta temática. Se ha propuesto en (D'Amato et al 2016) una plataforma distribuida que permite administrar y compartir, a través de la Web, un conjunto de cámaras IP con procesamiento embebido, en forma pública o privada. Esta idea surge como respuesta a la creciente demanda social en materia de seguridad y en sintonía con el aumento constante en la instalación de cámaras de seguridad que apuntan a la vía pública. A ello se suma una filosofía colaborativa que facilite el cumplimiento de una meta general: mayor seguridad social. Esta plataforma contempla la automatización de tareas para la detección y seguimiento de personas y objetos en diferentes cámaras de video, de manera integral. Su arquitectura abierta y distribuida contempla el uso de cámaras con procesamiento in situ para detectar rápidamente la información de interés en el mismo lugar en que sucede. Solo aquel evento importante será transmitido a una central, lo que permite escalar el sistema para soportar muchas cámaras sin afectar el rendimiento global.

En este trabajo, se presenta un análisis de cómo gestionar la información provistas por dichas cámaras a fin de crear servicios de información de distinto tipo, que permitan el accionar, tanto para el gobierno como la sociedad. El trabajo se divide en diferentes secciones. Primero se presentan algunos antecedentes tecnológicos y también se caracteriza lo que hace a una ciudad inteligente. A continuación, se proponen de servicios que pueden crearse a partir de estos datos y una de arquitectura integral que abarque a los diferentes actores. A continuación, se realiza un estudio de las necesidades de infraestructura de procesamiento requeridos para generar dichos servicios, a partir de la estimación de diferentes métricas en base a casos reales, considerando tiempos de respuesta y precisión de la información. Por último, se exponen las conclusiones y los trabajos para a futuro.

2. Estado del arte

Debido al volumen de información que se generan por video-vigilancia y de los sensores desplegados en una ciudad, se encuentran diferentes trabajos que abordan la problemática, ya sea por los algoritmos requeridos para el procesamiento como para la arquitectura que requiere un sistema de este tipo.

Desde el punto de vista algorítmico, en trabajos como el de *Serby et al. (2004)*, se detallan algunos de los métodos más comunes para la detección y seguimiento de objetos. *Hall et al. (2010)* realizan una comparación interesante entre los distintos detectores, en donde se concluye que la combinación de los mismos puede ser muy útil para disminuir la tasa de falsos positivos sin comprometer el análisis en tiempo real. Más recientemente, en *Ojha y Sakhare (2015)* se revisan las estrategias generales para la clasificación y el seguimiento de personas en video y en *Kim et al. (2011)* se plantea un método robusto en tiempo real, apto para el seguimiento de objetos a partir de predecir su velocidad y dirección. Además, en *Olivieri et al. (2012)* se describe un método denominado Motion Vector Flow Instance (MVFI), capaz de detectar actividades humanas complejas como “caminar” o “correr” con cámaras estáticas. Considerando las soluciones integrales existentes, en *Gdanks (2011)* se hace una revisión de 18 sistemas de videovigilancia comerciales y un resumen de los algoritmos más utilizados en etapas de detección, seguimiento y clasificación de distintos objetos.

Desde el punto de vista de la arquitectura, *Lin et al. (2011)* y *Hassan et al. (2015)* proponen utilizar la nube para resolver la escalabilidad, aunque están orientados únicamente al procesamiento de video, descartando la integración de otros tipos de señales. Pensando en ciudades digitales y sumado al bajo costo de un procesador embebido, es natural que se considere distribuir el análisis de imágenes. Incluso ya existen algunas soluciones comerciales, propuestas por NVidia que utilizan redes neuronales en el mismo hardware. En el trabajo (*D’Amato, et al 2018*), se ha propuesto una solución con múltiples dispositivos conectados que realizan análisis en el lugar, y que reportan únicamente cuando se detecta una situación anormal.

2.1. Hacia la conversión en una ciudad inteligente

Con el objetivo de poder realizar un plan de transformación de las actuales ciudades, en “ciudades inteligentes”, en (*ONTSI 2015*) se realizó un estudio de 200 municipios, a fin de entender el estado actual en dicho país. Como resultado del estudio, se llegó a un diagnóstico general de evolución de dichas ciudades, y se propuso un paquete de herramientas para la transformación digital. Principalmente, el resultado fue el diseño de una cadena de valor para una Ciudad Inteligente, que describe las diferentes actividades y fases a realizar en la transformación de una ciudad. La cadena de valor ha permitido realizar una caracterización de los agentes participantes en dichas Ciudades, distinguiendo entre empresas de distintos rubros tecnológicos, desde infraestructura digital hasta proveedores de plataformas y profesionales de información. Se ha identificado, el rol importante que tienen dichas empresas y en las cuales actualmente contribuyen con el ciudadano en brindarle asistencia digital.

En este contexto, el estado debe poder conjugar las capacidades disponibles y el aprovechamiento propio de las tecnologías y recursos disponibles, pero al mismo

tiempo, permitir que estas empresas puedan brindar nuevos servicios a la ciudadanía. (que no hacen al objetivo/misión de un municipio), preservando principios individuales del ciudadano y bajo una metodología de adopción transparente. La posibilidad de poder ofrecer los nuevos servicios requiere contar con infraestructura digital suficiente para poder crecer en cantidad de usuarios, en calidad de información y en menores tiempo de respuesta. Hoy ciertamente, el control de acceso a la información es potestad del gobierno, a través de sus diferentes dependencias. Por otro lado, esta concentración de responsabilidad hace difícil poder responder a los nuevos desafíos listados. Para lograr un sistema sustentable desde el punto de vista económico y escalable en el tiempo, se propone la idea de un modelo de aprovechamiento de los datos público-privado, asegurando equidad de acceso a la información y seguridad y privacidad de los datos, evitando el mal uso de la información disponible. Con ese fin, se han identificado algunas condiciones que se deben cumplir:

- Acceso controlado a los recursos públicos a partir del registro de “proveedores”
- Pre-procesamiento de los datos, eliminando información que permita una identificación de un individuo o vehículo
- Mecanismo transparente de beneficios del uso de los datos
- Control del flujo de la información interno al gobierno
- Compromiso de beneficio social
- Involucramiento de la sociedad, aportando datos propios

Las empresas, en este contexto, juegan un rol clave, ocupando aquellos espacios disponibles, sin desviar al gobierno de su rol y asegurando que los servicios lleguen a los reales interesados, haciendo toda actividad de gestión comercial propia de su objetivo. En este esquema, también debe idearse un esquema de economía, que beneficie a todas las partes, aunque escapa del alcance de este trabajo. Por último, en base a estos principios, es que se diseña la propuesta de sistema mostrado en la siguiente sección.

3. Identificación y caracterización de servicios digitales

La cantidad de información que generan los sistemas de video-vigilancia es enorme, ya que a partir de las imágenes es posible identificar y caracterizar el comportamiento de la población a partir de inferencias en los puntos de ubicación de las cámaras de video. En tal sentido, esta información permite describir principalmente el tránsito vehicular y peatonal, y detectar toda aquella anomalía en dicho comportamiento. A fin de identificar este comportamiento, es necesario recabar información espacio-temporal de manera continua y poder idear servicios basados en la información que sirvan a los diferentes actores de las ciudades.

Conociendo las capacidades que tiene un sistema de video-seguridad y sumado a los usos que realiza un municipio tipo de dicho tipo de sistema, se propone la caracterización de los servicios, en dos (2) tipos.

- **Para el gobierno:** Son aquellos servicios que competen al gobierno por su rol y que a su vez tienen el objetivo de contribuir a la planificación general (por ej., instalación de un semáforo en una esquina)

- **Para la ciudadanía:** son aquellos servicios que sirven a fines o intereses individuales o grupales, para una mejora o ayuda particular y ocasional

Por otro, en función a la urgencia de la disponibilidad de la información, se proponen también dos (2) sub-tipos:

- **Tiempo real:** eventos que deben ser atendidos en el momento que han sucedido
- **Tiempo diferido:** indicadores que no requieren ser atendidas, sino que sirven con fines estadísticos en las decisiones.

En base a esta primera caracterización, se han identificado diferentes servicios que es posible desarrollar, tal como se muestra en la Tabla 1, dentro de alguna de las dimensiones que definen una ciudad digital:

Destino	Tipo	Requerimiento de respuesta
Servicios para el gobierno	Seguridad	Tiempo real
	Cantidad de vehículos de gran porte	Tiempo diferido
	Eficiencia de los semáforos	Tiempo diferido
	Detección de camiones	Tiempo real
	Concentración de peatones	Tiempo real
	Comportamiento de peatones al cruzar	Tiempo diferido
	Comportamiento del conductor frente al peatón	Tiempo diferido
Servicios para el ciudadano	Detección de estacionamientos libres	Tiempo real
	Detección de animales sueltos	Tiempo real
	Detección de calles averiadas	Tiempo real
	Congestión media por hora	Tiempo diferido

Tabla 1 – Servicios identificados por destinatario y por requerimientos de respuesta.

3.1. Propuesta de arquitectura del sistema y de organización de la información

A fin de poder brindar algunos de los diferentes servicios listados anteriormente, es necesario poder determinar la información que se va a obtener en cada cámara. En particular, a partir del análisis únicamente de la imagen, se obtienen dos tipos de datos, propios de la escena (edificaciones, árboles, iluminación) y propios de los objetos en la escena en un momento dado. En este trabajo nos centraremos en el segundo tipo. En la identificación de objetos en cada momento de tiempo, es posible obtener los siguientes datos individualizados:

- La clase identificada (vehículo, persona, animal)
- Características visuales (forma, color)
- El desplazamiento
- El tiempo de presencia frente a la cámara

En base a estos indicadores básicos, es posible crear otros indicadores contextualizados de acuerdo al servicio que se quiera brindar. Si se desea aplicar a movilidad urbana, es posible crear indicadores tal como *Flujo vehicular por hora*, *Dirección del movimiento* o incluso *la velocidad del desplazamiento*.

La plataforma propuesta contempla la integración de cámaras IP, junto con herramientas de análisis automático y un gestor de eventos, preservando y controlando el acceso a los datos. Esta funcionalidad se agrupa por capas, una primera capa de gestión de los recursos tecnológicos que son gestionados en principio por el gobierno y luego cada una de las capas de transformación de la información. En la Figura 2, se muestran los diferentes componentes de la arquitectura. Esta arquitectura fue propuesta originalmente en *D'amato et al., (2016)* y ahora se reformula para soportar un nuevo esquema de acceso a la información, basada en servicios.

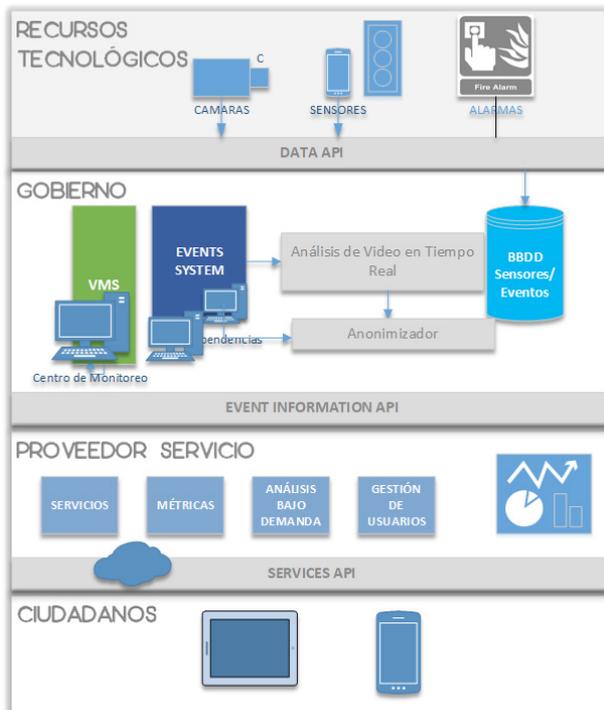


Figura 2 – Arquitectura en capas de los servicios provistos

En esta nueva arquitectura, todas las cámaras y demás recursos, como sensores o alarmas, se acceden a través de una API, la cual de momento es de acceso restringido al gobierno. Esta situación estratégica de punto intermedio le permite ofrecer diversas funcionalidades: control de acceso, registro del tráfico, restricción a determinados tipos de recursos, mejora de rendimiento, anonimato de la comunicación, etc. Las cámaras y sensores generan nuevos flujos de datos que ingresan a la plataforma a través de aun

API de datos, que luego es procesada por gestores de Videos o plataformas de Sistemas de Eventos.

A continuación, se realiza un análisis de dichos datos, principalmente los algoritmos de clasificación y caracterización de objetos pueden ser de múltiples tipos y con diferentes objetivos. Desde algoritmos más simples de clasificación por la forma y tamaño, hasta otros específicos como detectores de patentes y rostros. Cada algoritmo tiene una complejidad computacional diferente y utilizan diferentes características de las imágenes observadas. En cualquiera de los casos, lo que se pretende es poder describir un objeto con el mayor detalle posible. Al finalizar el proceso, en caso que sea necesario, se propone realizar una etapa de anonimización (por ejemplo, ocultando patentes o caras visibles de personas), se registran temporalmente estos datos, en caso que haga falta, y se ponen a disposición también a través de una API de Eventos, basada en tecnología REST. Entre ambas capas se gestiona el tipo de servicio, ya sea si es estadístico o en tiempo real. Los datos se intercambian en formato JSON, tal como se muestra en la Fig. 3.



Figura 3 – Componentes del servidor y de la vista en un caso de instanciación.

A través de esta API, los diferentes proveedores de servicios cuentan con la información de cada objeto detectado, con el momento y ubicación de su detección. En la capa del proveedor del servicio, se debe gestionar toda aquella funcionalidad que haga a este rol: generación de estadísticas, gestión de usuarios, visualización de los datos, datos históricos, registros de servicios o alertas, entre otros. También recae en esta capa, todo aquello que signifique interacción con el usuario a través de interfaces WEB o Aplicaciones móviles. La infraestructura para desplegar esta capa puede ser realizada con tecnología Cloud, pensando en una mayor escalabilidad y facilidad de conexión.

3.2. Generación de información en tiempo real y distribución del procesamiento

El primer paso para la detección de eventos es poder determinar si hubo o no movimiento a partir de una serie de imágenes. Este problema es análogo al problema de sustracción de fondo, que calcula la variación entre imágenes consecutivas (Shaikh et al. 2014). Los métodos varían en complejidad computacional y eficacia del resultado. Actualmente, en la plataforma se utilizan variantes de VIBE (Bouwman et al., 2014), ya que este algoritmo permite descartar movimiento continuo, como por ejemplo el movimiento de

las hojas de los árboles. Cuando las cámaras son exteriores, se suceden otros factores que además varían dependiendo del momento del día, la nubosidad y la época del año..

Existen múltiples algoritmos especialmente desarrollados para la detección de personas y vehículos. Muchos de estos algoritmos son de código abierto y otros a través del uso de librerías de terceros. Los algoritmos varían en velocidad, precisión, y tipo de datos necesarios. Algunos utilizan una imagen del video, otros una secuencia de imágenes, o en algunos casos sólo alguna información como el movimiento realizado. El resultado es una caracterización más detallada del objeto, que permite una mejor clasificación. Para la clasificación de estos objetos, utilizamos dos (2) configuraciones: una utilizando solo redes convolucionales basadas en el trabajo de *Redmon, Farhadi (2018)*, y otro una combinación de object-tracking con estas redes convolucionales (D'Amato et al 2018). Por cada nuevo objeto detectado en tiempo real, se genera un mensaje asociado a un tópico. A este se le asocia una imagen que lo representa y el tiempo de captura, en caso de que se deba acceder al flujo de video para ubicarlo. Este tópico se almacena en la base de datos y se pone disponible a través de la API con un identificador único. Las salidas de estos algoritmos se almacenan y transmite en formato JSON. Cada mensaje de respuesta viene asociado con información del algoritmo que lo realizó, tal como el tipo de análisis, el tiempo de respuesta y la confiabilidad del resultado.

4. Análisis y resultados

En base a esta arquitectura, actualmente se están desarrollando dos (2) aplicaciones. El cliente de una de estas aplicaciones son las fuerzas policiales, el poder judicial y otras instituciones gubernamentales externas al Municipio que utilizan la información como herramienta de diagnóstico y búsqueda eficiente. También, se está desarrollando otra aplicación, con la misma base, para generar estimadores de movilidad urbana, que es utilizada por los organismos encargados del planeamiento y organización del tránsito urbano. Las aplicaciones se desarrollan utilizando tecnología WEB, tal como AngularJS® para la visualización, siendo multi-plataforma y adaptable a los diferentes dispositivos. Dicha aplicación consume una API Rest propuesta por este equipo, para obtener los datos. La aplicación cliente cuenta con interfaces gráficas para acceder, consultar datos históricos y herramientas gráficas para manipular las búsquedas. De momento, el acceso a los datos se encuentra centralizado en el gobierno. Para poder estimar el costo de contar infraestructura (equipamientos de procesamiento), para ofrecer estos servicios realizamos una evaluación de la capacidad de proceso que tiene actualmente un servidor. A continuación, generamos diferentes ejemplos de servicios a partir de los datos hoy disponibles en la plataforma.

4.1. Evaluación del esfuerzo para el procesamiento

Con el fin de analizar la factibilidad de la distribución del procesamiento, se realizaron diversas pruebas con diferentes equipamientos:

- Conf. 1: instancias virtuales en servidor Xeon E5 20 a 1.7GHZ.
- Conf. 2: PC Intel i5-7300HQ a 3.0 GHZ con placa gráfica GTX1050.
- Conf. 3: PC Intel i7-7800 a 3.5 GHZ con placa gráfica GTX1060

En esta prueba, se buscó evaluar el impacto de diferentes configuraciones de hardware y dos configuraciones de redes neuronales Convolucionales, con mayor velocidad, pero menor precisión y viceversa. En la tabla 1 se muestran los resultados obtenidos, aplicados a 1000 frames consecutivos de un video. Para la clasificación, se utilizaron dos algoritmos: uno propio basado en tracking y sustracción de fondo con clasificación usando una red neuronal y una adaptación propia de la biblioteca YOLO para la clasificación de los objetos usando OpenCL. Se utilizaron cámaras de alta resolución (1920x1080).

Configuración	FPS CNN Yolo	FPS Vibe +CNN
High PC – no GPU	1,2	43
Middle PC – middle GPU	2,9	75
High PC – High GPU	5,4	124

Tabla 2 – Cómputo de la cantidad de procesamiento por segundo

A su vez, se midió la tasa de detección de objetos de nuestra red con tracking (Vibe + CNN), obteniendo una tasa de detección del 85.5%, mientras que la red Yolo fue del 97.5% aproximadamente. En este segundo caso, la tasa de acierto es muy alta y es efectiva cuando existen gran cantidad de objetos de manera simultánea en la escena. Por otro lado, nuestro algoritmo sirve cuando la cantidad de objetos en escena es relativamente baja. La cantidad de frames por segundo que puede procesar un equipo indica a la infraestructura necesaria para el procesamiento. Si se desea realizar procesamiento en tiempo real, se requieren por lo menos 20 FPS. Por lo que una PC de alto desempeño (configuración 3) puede llegar a procesar aproximadamente 6 cámaras de manera simultánea, mientras que un servidor virtualizado sólo puede procesar hasta 2 cámaras por cada instancia. Se requiere una combinación de técnicas para poder dar servicios en tiempo real, o sacrificar calidad en la detección.

4.2. Generación de servicios de movilidad

Utilizando la configuración de más precisión, se crearon diferentes indicadores que sirven para caracterizar la ubicación de una cámara. Se propusieron dos situaciones de análisis diferente, y se generaron diferentes indicadores que sirven en los diferentes servicios. Los casos propuestos son:

- Caso 1: calle de alto tránsito en el centro de la ciudad
- Caso 2: ruta a la salida de la ciudad.

Se generaron los siguientes indicadores estadísticos, a fin de validar que permiten la caracterización de diferentes ubicaciones y comportamientos. Primero, una vista de mapa de calor, lo que tiene un valor cualitativo (ver figura 3) y el equivalente por hora.

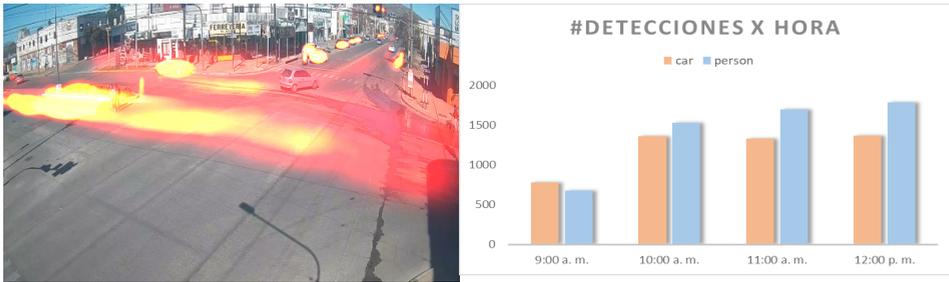


Figura 3 – (izq) Mapas de calor (der) Tránsito discriminado por hora por clase

Al mismo tiempo, representamos la concentración de objetos por su clase, en los dos casos. Las clases principales son: auto, persona, animales, bus, camión, motocicleta y otros. Las proporciones de cada clase se observan en la Figura 4.

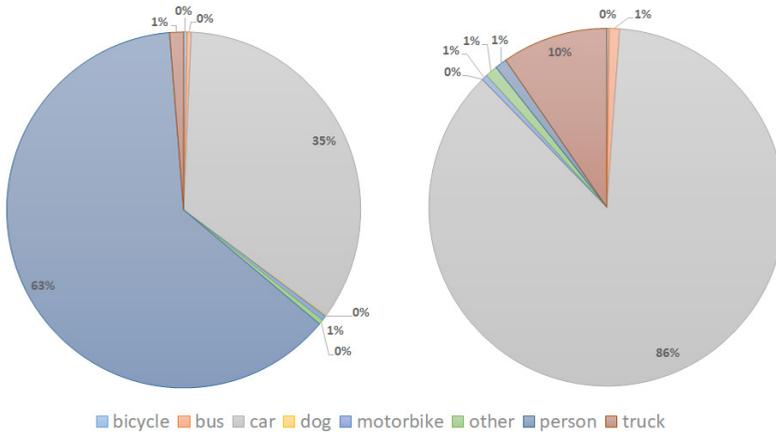


Figura 4 – Concentración de objetos por su clase (izq) Caso 1 (derecha) Caso 2.

En dichas representaciones, es notorio la concentración de las diferentes clases, lo que ayuda a caracterizar una ubicación. En el **caso 1**, se observa una alta concentración de personas, y una menor del resto de los vehículos, pero aparecen casos de animales sueltos. Por otro lado, en el **caso de la ubicación 2**, hay una mayor concentración de los diferentes vehículos, con menor o nula situación de personas o animales.

5. Conclusiones

En este trabajo se presentó la idea de generar servicios basados en información para ciudades inteligentes a partir del uso de datos extraídos en video-cámaras. Para eso, se ideó una arquitectura que integra los diferentes integrantes: empresa, gobierno y ciudadanía y una plataforma distribuida para la gestión de múltiples cámaras y sensores.

La plataforma contempla la definición de una API que permita integrar aplicaciones externas.

Desde el punto de vista técnico, se mostró que, para obtener una plataforma funcional, se requiere combinar algoritmos eficientes. La configuración final de la plataforma depende de qué se priorice: o bien reducir el tiempo de respuesta o bien la confianza en los resultados. El manejo de esta información tiene un impacto en el esfuerzo computacional, y a su vez, en la infraestructura que se debe gestionar. Esta información puede ayudar en la ubicación o configuración de semáforos inteligentes, para definir políticas de bandas horarias para zonas de peatones, entre otras. A su vez, la tipificación de concentración de tipos de vehículos sirve con fines comerciales. Por otro lado, existen cuestiones a profundizar, tal como los mecanismos para acceder estos datos, de manera de asegurar la privacidad de los datos, y evitar su utilización con fines delictivos. Como trabajos futuros, se pretende continuar, ampliando las posibilidades y combinando recursos de diferente tipo, tal como la integración con semáforos o iluminación.

Agradecimientos

Este proyecto ha utilizado datos obtenidos en el marco de colaboración Univ. del Centro (UNCPBA) y Municipio de Tandil, Argentina. A su vez, ha recibido subsidios de la CICPBA, y el subsidio joven investigador 2016, CONICET (Argentina).

Referencias

- Agrawal, H.; Mathialagan, C. S.; Goyal, Y.; Chavali, N.; Banik, P.; Mohapatra, (2015). "CloudCV: Large Scale Distributed Computer Vision as a Cloud Service". En Hua, G. y Hua, X. *Mobile Cloud Visual M.C.* Springer, pp. 265-290.
- Bouwmans, D.; Porikli, F.; Hoferlin, B. y Vacavant, A. (2014). *Background Modeling and Foreground Detection for Video Surveillance*. Londres: Chapman and Hall.
- Carrión, F.; Pontón, C. y Armijos, B. (2009). *120 Estrategias y 36 experiencias de seguridad ciudadana*. Quito: Flacso.
- D'Amato, J.P., Negri, P. Lotito, P. (2011) *Estimating the Queue Length to Optimize the Green Time for an Urban Traffic Control System*. III Congreso MACI, Argentina.
- D'Amato J. P.; Domínguez, L.; Pérez A. y Rubiales A. (2016). "Plataforma abierta de gestión de cámaras IP y aplicaciones móviles para la seguridad civil ciudadana". *RISTI- Ibérica de Sistemas e Tec. Inf.*, vol. 20, pp. 48-61.
- D'Amato J. P.; Domínguez, L.; Pérez A., Rubiales A., R. Barbuzza, F. Stramana (2018) *Video Analytics on a Mixed Network of Robust Cameras with Processing Capabilities*", *Advances in Intelligent Systems and Computing book series (AISC, volume 745)*
- Dammert, L.; Salazar, F.; Montt, C. y González, P. (2010). *Crimen e inseguridad: indicadores para las Américas*. Washington DC: FLACSO-Chile/BID.

- Gdanks, Ku (2011). "Deliverable 2.1. Review of existing smart video surveillance systems capable of being integrated with ADDPRIV project". Recuperado de <http://www.addpriv.eu/uploads/.../GDANSK_Scoreboard_R11.pdf>.
- Hall, D.; Nascimento, J. y Ribeiro, P. (2005). "Comparison of Target Detection Algorithms Using Adaptive Background Models". En: 2005 IEEE Int. W. on Visual S. and Performance Evaluation of Tracking and S.. IEEE, pp 113–120.
- Hassan, M. M.; Hossain, M. A.; Abdullah-Al-Wadud, M.; Al-Mudaihesh, T.; Alyahya, S. y Alghamdi, A. (2015). "A Scalable and Elastic Cloud-Assisted Publish/Subscribe Model for IPTV Video Surveillance System". Cluster Computing, vol. 18, pp. 1539-1548.
- Ixmattlahua, S. D.; Raygoza, R. O.; Romero, O.; Uribe, F. y Vargas, E. J. (2015). "Metrópoli digital: una plataforma web para la inclusión integral de las PyMES, soc. y gob. en el uso de las tecnologías de la información en la región de las altas montañas del estado de Veracruz, México". RISTI, n.º 3, pp. 43-54.
- Hanxi Li, Yi Li, Fatih Murat Porikli (2016) DeepTrack: Learning Discriminative Feature Representations Online for Robust Visual Tracking. IEEE transactions on image processing , DOI:10.1109/TIP.2015.2510583
- Kessler, G. (2009). El sentimiento de inseguridad: sociología del temor al delito. Buenos Aires: Siglo XXI.
- Kim, J. S.; Yeom, D. H. y Joo, Y. H. (2011). "Fast and Robust Algorithm of Tracking Multiple Moving Objects for Intelligent Video Surveillance Systems". IEEE Transactions on Consumer Electronics, vol. 57, n.º 3, pp. 1165-1170.
- Mahjoub, M.; Mdhaffar, A.; Halima R. B. y Jmaiel M. (2011). "A Comparative Study of the Current Cloud Computing Technologies and Offers". En 2011 First Int. Symposium on Network Cloud Computing and Apps. IEEE, pp.131-134.
- Ojha, S. y Sakhare, S. (2015). "Image processing techniques for object tracking in video surveillance - A survey". En: 2015 International Conference on Pervasive Computing (ICPC), pp. 1-6. doi: 10.1109/PERVASIVE.2015.7087180
- Olivieri, D. N.; Conde, I. G. y Sobrino, X. A. V. (2012). "Eigenspace-based Fall Detection and Activity Recognition from Motion Templates and Machine Learning". Expert Systems with Applications, vol. 39, n.º 5, pp. 5935-5945.
- ONTSI (2015a). Estudio y Guía metodológica sobre Ciudades Inteligentes. Recuperado de: [http://www.ontsi.red.es/ontsi/es/estudios-informes/..](http://www.ontsi.red.es/ontsi/es/estudios-informes/)
- Redmon, J. Farhadi, A. (2018) YOLOv3: An Incremental Improvement, ArXiv, <https://arxiv.org/pdf/1804.02767>
- Shaikh, S. H.; Saeed, K. y Chaki, N. (2014). Moving Object Detection Using Background Subtraction. Berlín: Springer International Publishing.
- Smart city framework – Guide to establishing strategies for smart cities and communities. BSI Standards Publications. PAS 181:2014. Feb. 2014