Computer Vision

External note

April 2023

Use cases and typical infrastructures studies with Orange

Executive summary

Computer Vision brings added value to multiple activities, in different economic domains

Computer Vision, by acquiring digital images and generating information from the understanding of them, provides added value to numerous economic sectors – industries, cities, retail, territories – and activities – security, quality insurance, measurement, marketing knowledge.

Orange takes part of the Computer Vision ecosystem

Orange business units, and especially Orange Business, take part of this growth domain as infrastructure provider, integrator, and Computer Vision solution provider.

Computer Vision use cases can leverage on Cloud, Edge, 5G, WiFi, to target their needs

The variety of uses and rolled-out conditions brings different needs and constraints, that a single cloud-based infrastructure cannot all meet. 5G and Edge provide new capabilities to Computer Vision, especially in industries, cities, retail and territories:

	Network(s)	Computer Vision treatment localization
Industries		
Quality Insurance in industry	Ethernet → 5G	Near edge → on-premise
Worker safety in factories	5G	On-premise
Video monitoring by drone	5G	Cloud or on-premise
Cities, retail, territories		
Augmented video protection in cities	Ethernet, WiFi	Private cloud
Security in public transportation	4/5G, Ethernet	Private cloud
Customer behavior analysis in retail	WiFi, Ethernet	On-premise
Forest fire detection	4/5G	Edge

Computer Vision brings added value to multiple activities, in different economic domains

Computer Vision has made, during the recent years, spectacular progresses, thanks to technological advancements and opensource contribution in software and datasets:

- **Deep Learning**[1], in conjunction with older image manipulations techniques, has highly improved the capacity to identify information in scenes, and its accuracy.
- **Capacities and prices of cameras**[2]**, networks**[3]**, compute capacities**, either in Cloud or at the Edge, and **the capacity for non-technical people to define and make recognition model evolve**, has permit to use this technology on a large scale.

Computer Vision development has also been catalyzed by the increase of the stakes it can answer:

- **Automatic and non-human detection of fault**, to answer to competition in industry, and necessity of quality at a reasonable cost.
- **Security-related human behavior detection**, in many situations, and reinforced by COVID crisis, and China particular approach of people control.
- **Environment recognition and measurement**, which is mandatory for new and future usages such as semi/autonomous vehicles.

This trend is seen to be continued and amplified in the next years especially thanks to new AI generation.

Computer Vision applications are used in different activity domains, and especially in:

- **Industry**, with Quality Insurance, but also positioning and guidance for Robotics and Automated Guided Vehicle (AGV)[4], identification for worker safety.
- **Smart City and building**, with identification for safety and infrastructure planification.
- **Retail**, with identification for customer behavior knowledge.
- **Transport**, with (semi)-autonomous vehicles.
- **Health**, for automatic medical image analysis.
- And also in **B2C, for Home** monitoring and security.

[1] Artificial intelligence-based algorithms

[2] Higher resolutions, higher performance interfaces, embedded vision & edge computing in camera, 3D image acquisition, image data beyond the visual spectrum, sensor fusion, to field of view and depth of field, sensitivity, dynamic range, frame rate and resolution

[3] cable length, light conditions, latency and jitter, throughputs

[4] AGV uses cases defined in internal 5G ACIA (Alliance for Connected Industries and Automation) report and public white paper named 5G-ACIA Report WI 055 5G Industrial Edge Use cases and Requirements.

Orange takes part of the Computer Vision ecosystem

Orange Business proposes different offers to help and take part of the Computer Vision market growth:

- Orange Business proposes first its build expertise and run management of a **Cloud + Edge infrastructure**, with dedicated AI and Machine Learning modules, and **connectivities** (4/5G, WiFi), either public, private or hybrid.
- Orange Business also proposes its own **Computer Vision solution**, based on its platform, AI algorithms and expertise to adapt if needed these treatments to use case specificities.
- Based on these Orange bricks, and also on market solutions, Orange Business is a **global integrator and operator of the solution**, including end to end security assurance.

Orange is part of an ecosystem of Computer Vision providers, with competitors, or coopetitors, these acting as partners, either using Orange infrastructure or completing Orange Business Computer Vision offer, for specific use cases.

Hyperscalers are also present in this ecosystem, proposing to all services providers their hosting infrastructure, their deployment and management tools, and AI-oriented native services.

Based on market and domain survey, and confirmed by the experience of Orange Business and Orange Innovation teams, some Computer Vision use cases can be considered as both promising in terms of business, and representative in terms of constraints and needs.

We detail in this paper 3 use cases in industries:

- Quality Insurance, on a production line in industry, with customer experience and feedbacks in Poland.
- Worker safety in factories.
- Video monitoring by drone, with Arboreal Intellbird and Orange Spain experience.

And 4 use cases in cities, retail, territories:

- Augmented video protection in cities.
- Security in public transportation (bus, tramway, metro), with Cities experiences in France.
- Customer behavior analysis in retail, with partner experience in a clothing shop in France.
- Forest fire detection.

Quality Insurance, on production line in industry → Polish customer experience and feedbacks

As seen in the previous chapter, Quality Insurance is a major application domain of Computer Vision: this technology allows to compare automatically / without human intervention, at any step of a production line, the current status of the product with its expected status at this step. The model on which the comparison is done can evolve, with incremental, supervised, … machine learning.

Image processing systems measure and count products, calculate their weight or volume, and inspect goods at top speed with respect to their predefined characteristics. Furthermore, they automatically extract limited, but crucial, information from huge quantities of data, or they help experts interpreting images by filtering, optimizing, supplementing, or quickly retrieving and making them available.

In a nutshell, machine vision systems enable the user or the application software to make the best possible decision in a timely manner.

In that capacity machine vision systems are used in a variety of applications, including manufacturing, traffic monitoring, warehouse automation, security systems or medical applications.

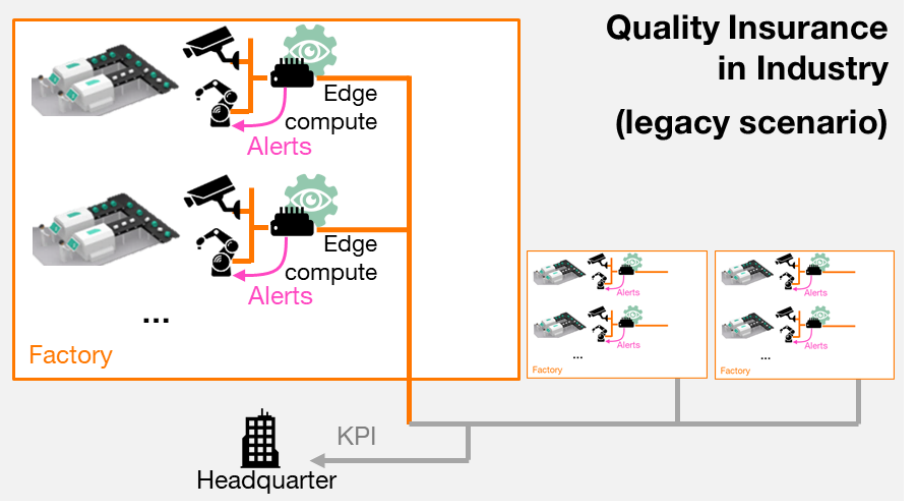
For sure, every situation is specific, but the customer case, where Orange has worked in Poland, allows to identify some needs and constraints of such a use case:

- **Number of cameras** usually start with a low basis, to answer to the first expected controls, but then can increase with the lines and the steps to be inspected → Polish customer case started with 12 cameras, for a global surface of factory of 6000 m2.

- **Quality of image** should be good and constant, since image details must be analyzed → In Polish customer case, each camera generates an upload speed of 10-12 Mb/s[5]. For this reason, Orange had to configure two 5G cells (5G is high download speed of 850 Mb/s but still low upload speed of 50-60 Mb/s – one carrier 3,6 GHz – 80 MHz block). All data traffic is sent over Orange core network and returned via fiber to the server room of the factory.
- **Localization of data** is mandatory: for security and resilience reasons, the video stream must be treated locally.
- **Treatment** is tailor-made and can be complex: hosting of the AI logic should be imagined more on a server, than on smaller compute capacity and/or the camera itself.
  ➔ Ex.: For both reasons, in Polish customer case, video stream is sent to a server, local to the factory (in its server room).
- Note that this use case does not bring **GDPR or acceptability issues**, since no human images are captured.

Quality Insurance: in-factory IA server, connected in 5G or WiFi6

- For security, isolation of data and latency, the use of a local treatment capacity is highly recommended. This capacity is used for IA complex treatments.
- In **legacy installations**, cameras and automatons used for feedback loop are wired connected locally: to leverage on this existing infrastructure, compute capacity can be installed locally at the edge, providing low latency. Edge capacities are managed in the factory, and send only indicators, for global consolidation.

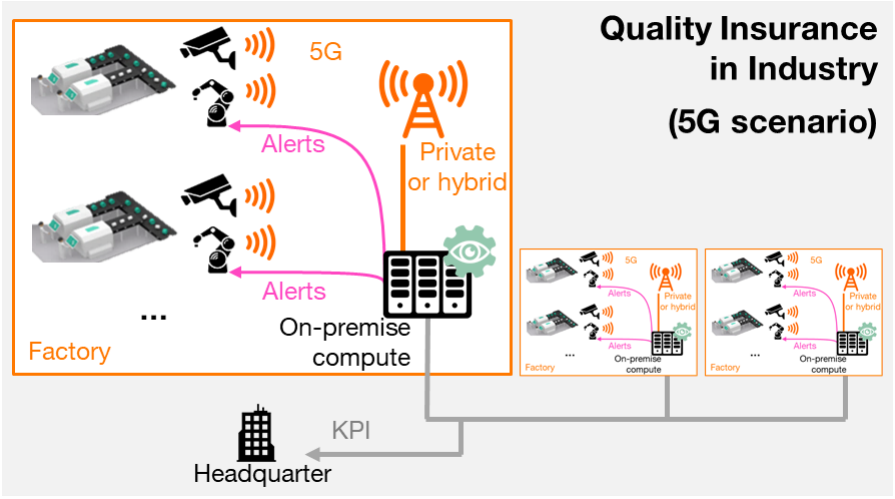


- With the evolution of devices in 5G context, a more centralized infrastructure can be imagined, based on 5G: native-5G cameras and automatons are linked to factory-centralized compute capacity. Low latency retroaction loop is provided by 5G. 3.6 GHz 5G network proposes a data

[5] In the document, different video formats are used. They refer to following definition and throughput:

4K	4096 x 2160 Pixel	From 10 000 to 35 000 kbit/s
Ultra HD	3840 × 2160 Pixel	From 10 000 to 35 000 kbit/s
Full HD	1080p, 1920 x 1080 Pixel	From 3 000 to 9 000 kbit/s
HD	720p, 1280 x 720 Pixel	From 1 500 to 6 000 kbit/s
SD	360p, 640 x 360 Pixel	From 400 to 1 500 kbit/s

transfer at a more than 750 Mbps speed, with easiest deployment than Ethernet wiring. Beyond Computer Vision, 5G offers an open infrastructure, for other use cases in industry. Note that WiFi6 could be another option (but not WiFi5, mainly for bandwidth and latency reasons).



- In the Polish customer case, cameras were not natively 5G, but connected in 5G by routers. This was due to lake of availability of 5G cameras on the market, which should evolve in the future.

Worker safety in factories

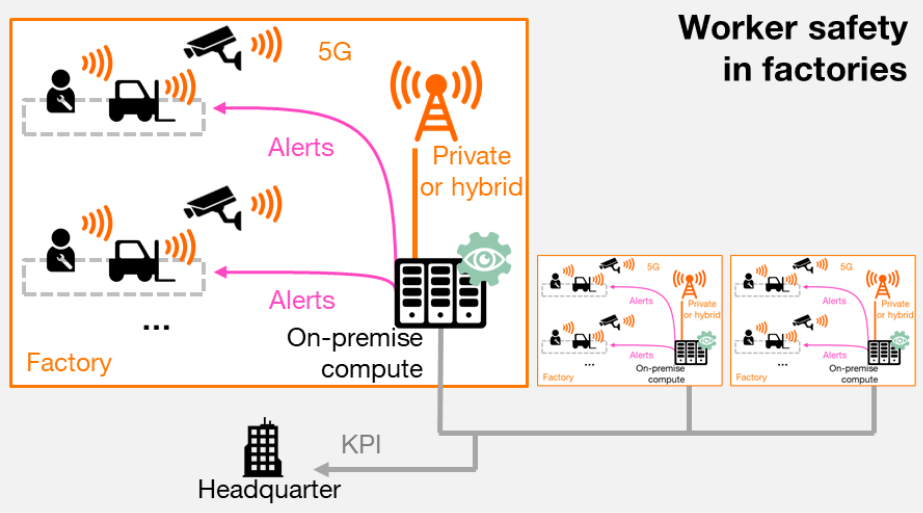
Identification capacities of Computer Vision are used to increase worker safety, indoor and outdoor: detecting position of people in at risk areas, verifying the use of security equipments such as helmet, glasses, gloves, preventing potential collisions between vehicles (ex. AVG – Automated Guided Vehicle) and persons or vehicles, alerting in case of fall of persons, doing prevention by enlightening (heat map) at risk zones.

- **Number of cameras** depends on the size of the area to be covered, and can be considered as installed specifically → 5-10 cameras per factory
- **Quality of image** should be good and constant, since image details must be analyzed. full HD to 4K cameras should be recommended.
- **Localization of data** is mandatory, first for latency reasons, but also for security and resilience reasons, the video stream must be treated locally.
- **Treatment** is tailor-made and can be complex: hosting of the AI logic should be imagined more on a server, than on smaller compute capacity and/or the camera itself.

Worker safety in factories

While radio protocol is mandatory, to equip persons and vehicles in mobility, 5G use and in-factory centralization of the compute provide key advantages:

- Possibility to scale treatment capacity, while increasing the number of cameras / areas to be secured.
- Avoid costly wiring in the factory, while leveraging on 5G good coverage and low latency.



Video monitoring by drone, with Arboreal Intellbird and Orange Spain experience.

Drones are more and more used to capture images of installations that cannot be accessed at low cost by other ways, or viewed with this from-the-sky angle. Computer Vision can be added to extract relevant information from images. Collaboration between video, drones and Computer Vision is used for surveillance, monitoring, inspection.

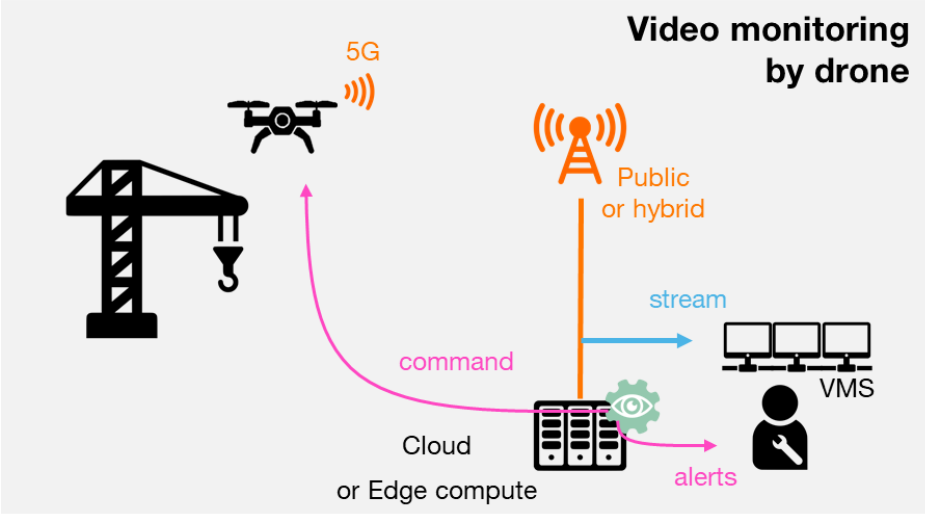
Orange Spain is working with Arboreal Intellbird, for instance for industrial inspection of electric lines.

- **Number of cameras**, equivalent to number of drones is today limited. But this number is seen to increase in the future, with more drones, doing everyday and simpler scenarios.
- **Quality of image** is key and full HD to 4K cameras should be recommended.
- **Localization of data** is not an issue. Since radio communication is seen as mandatory, treatment can be localized either in Cloud or in the premises. If retroaction is needed, either network or customer Edge will become the solution to reduce the latency of feedback loop.

Video monitoring by drone

Infrastructure proposed in the project is the following:

- 2K+ cameras.
- Cloud or edge treatment capacity



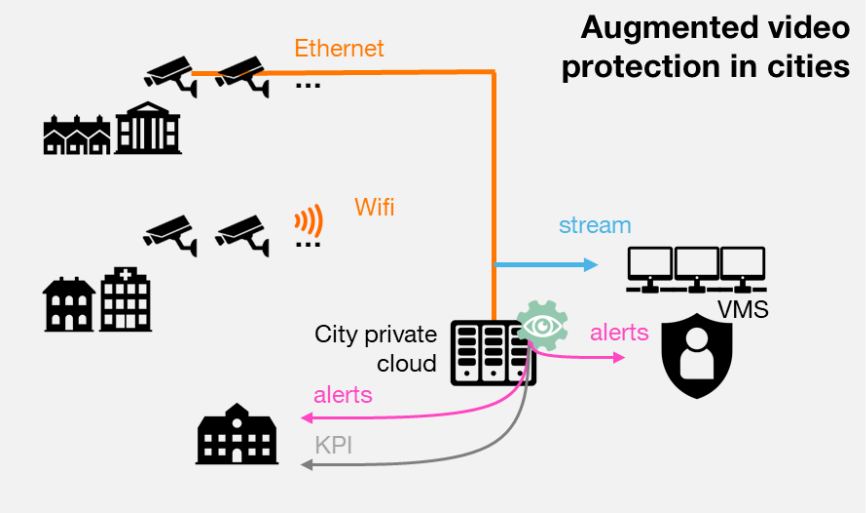
Augmented video protection in cities

Identification capacities of Computer Vision are also used in cities, to improve city operation and increase protection of citizens, coupled with video monitoring. The possible use cases, such as people counting, detection of crowd moves, weapons, gathering or assault, fall typically in "augmented camera" uses, identified by CNIL. With Paris Olympic Games, evolution of the law is to be followed closely.

- **Number of cameras:** for a middle range city, the number of cameras can reach a few hundreds. Orange had answered in the past to a major French city RFP, where 150 were proposed. Cameras are intended to be deployed mainly in the center, and around sensible places (schools, stadiums, …).
- **Quality of image**: for cities, video monitoring existing and renewed cameras should be used. These cameras are connected to the Video Monitoring System (VMS) of security forces. For cost reasons, and because needed AI treatments work well on this quality, 720P or 1080P cameras are sufficient.
- **Localization of data**: stream flow and extracted information is highly sensible and under GDPR constraints.
- **Computer Vision treatments** are various, to implement the multiple use cases targeted by the city.

Augmented video protection in cities

- Cameras are mainly connected through Ethernet wiring (and Powered Over Ethernet) and by WiFi, depending on the covered areas. Cameras on public 4G/5G networks are not expected to be used, or only in few and rare exceptions, in the city, to avoid high cost of communication.
- Capacity of treatment is centralized, in the city data center.
- Even if, in the same city, a security system is deployed in transport, both Computer Vision infrastructure are separated, since owners and operators are different (technical service of the city and transportation company services).



Security in public transportation → Cities experiences in Europe

Computer Vision is used for real-time video analysis, in public transportation (buses, tramway, metro, stations), for security and operational stakes: people counting, occupancy of disabled persons area, detection of weapons, abandoned luggage, fall of persons.

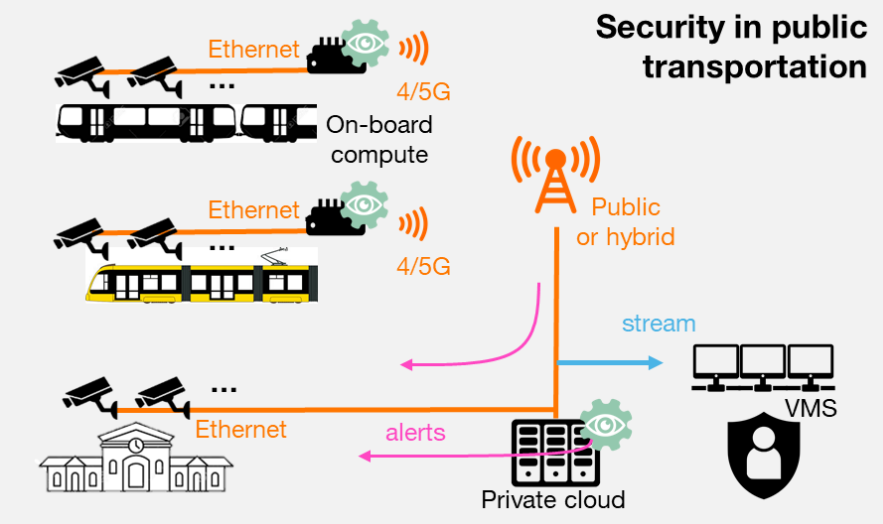
Based on different customer cases managed by Orange Business in cities in France and Europe, typical constrains and needs for this use case are the following:

- **Number of cameras**: every bus is usually equipped with 2-5 cameras, 8-10 for tramways and metro trains. With equipped stations, it can represent globally several hundreds and potentially thousands of cameras.
- **Quality of image** depends on existing or specific rollout, but classic configurations can involve cameras from HD Ready (720p), full HD to 4K (which generate an upload of 500kb-2Mb/s per camera). Note that cameras can be fixed or motorized and remotely pilotable (Pan Tilt Zoom – PTZ).
- **Mobility**: by construction, cameras can be either in mobility (buses, metro) or fixed (in station).
- **Connection to Video Monitoring System (VMS)**: since cameras first used for monitoring by humans, then connected to IA, both infrastructures should be connected.
- **Security and GDPR**: for security reasons, IT infrastructure of transport facilities is isolated from the Internet. Computer Vision infrastructure is part of it. Additionally, when available, video stream is only accessible by accredited people (security forces). For GDPR reasons, images should be

blurred/anonymized when used outside of local network, for build and test. Only information and alerts can be shared centrally.

Computer Vision in transport: a typical Edge/on-board + private Cloud infrastructure, on hybrid networks

- For security reasons, Edge computing is preferred to treat in-mobility images, as close as possible to the cameras. Localizing treatments in an edge appliance (and not directly in the camera) allows to be able to compute images coming from several cameras. This CPU/GPU capacity is deployed by bus and metro/tramway train. Remote streaming remains possible, through the VMS system, in specific situations (event, dispelling of doubt).
- For fix cameras, stream is usually already used by supervision. Treatment of these images can then be done centrally.
- Transportation IT context brings to use existing private Cloud, localized on-premise in the Control Center.
- For similar isolation reasons, communications are done using private Ethernet for fix cameras, 4G (and tomorrow 5G) private network for in-mobility underground situation, and public or hybrid 4G/5G network outside. This network is connected to VMS, used by security forces.



Customer behavior analysis in retail → Clothing shop partner experience

Computer Vision can be used for multiple purposes in retail, both for operational and customer knowledge stakes: people counting including or not gender/age criteria, waiting time in queue, hot zoning, behavior in front of products, …

- **Number of cameras** depends mainly on the size of the store, but with an average of 2-3 cameras by store.
- **Quality of image** is usually limited by Return of Investment, which is a high concern: HD Ready (720p) or full HD are preferred. To limit number and cost, 360° cameras are often used (with image adjustment).
- **Localization of data** is also driven by ROI and by existing capacity of communication connection (ADSL) of the store. Security of images and data is not seen as a major stake.

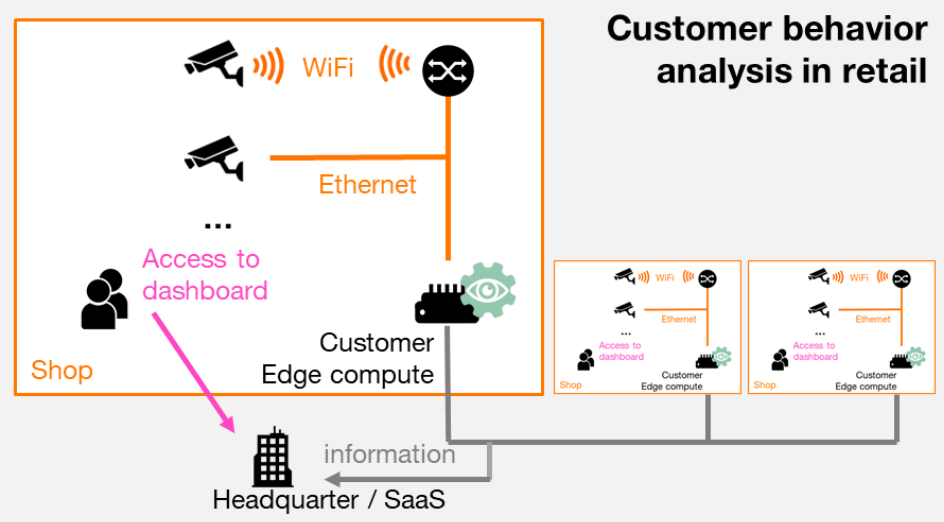
- **GDPR** must be a high concern since most of the use cases are based on human images capture. → Generation of statistics, and not individual behavior information has probably the main the way, for Orange partner, to reach CNIL acceptance.

Customer behavior analysis in retail: small local infrastructure for the best ROI

Infrastructure answer is highly piloted by costs:

- Considering the limited existing store communication connection (*), IA treatment is done locally, on a limited CPU/GPU capacity. Only calculated information and alerts are sent centrally, consolidated on Cloud/SaaS, and then accessed by people in the store.
- Cameras are connected through existing Ethernet (and are Powered over Ethernet), or in WiFi.

(*): Note that FTTH should not lead at short term to change to a central/cloud IA infrastructure, because of associated costs.



Forest fire detection

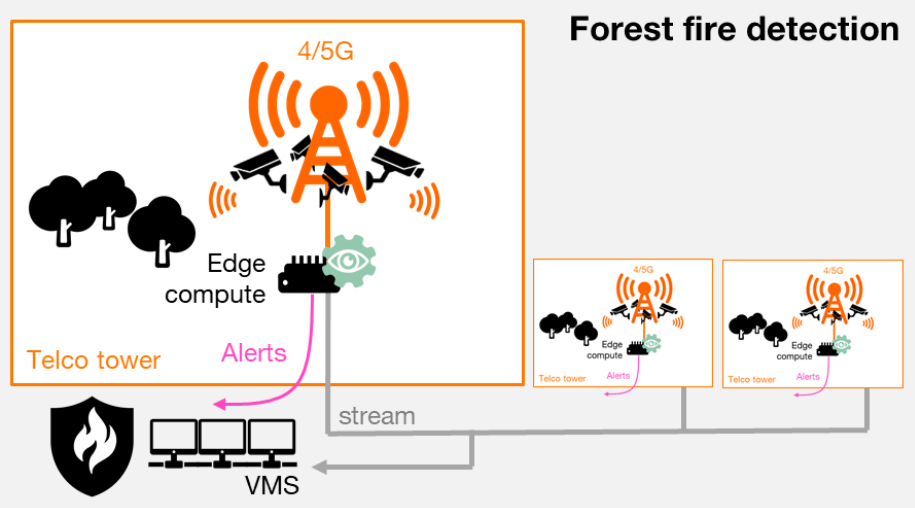
Computer vision can be used to detect smoke and fire, especially for the protection of forest.

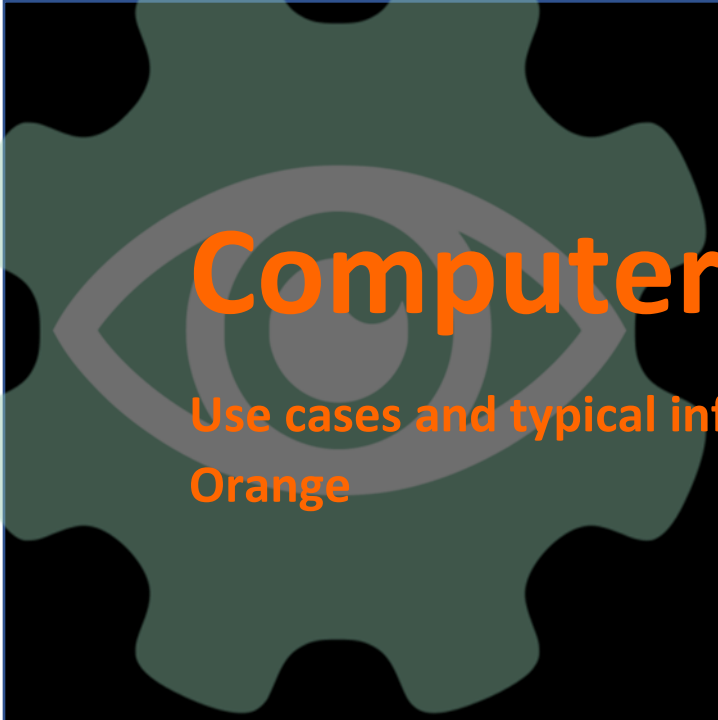
- **Number of cameras** depends on the area to be monitored.
- **Quality of image** is key and full HD to 4K cameras should be recommended. Camera would also gain to have also a thermic feature, in order to detect fire temperature.
- **Localization of data** is recommended, to limit the use of mobile bandwidth, with the possibility to ask for streaming video only on alert situations.

Forest fire detection

Infrastructure proposed in the CRISTEL project is the following:

- 4 full HD + thermic camera by Telco tower.
- Local edge treatment capacity





External note

April 2023

Computer Vision

Use cases and typical infrastructures studies with Orange