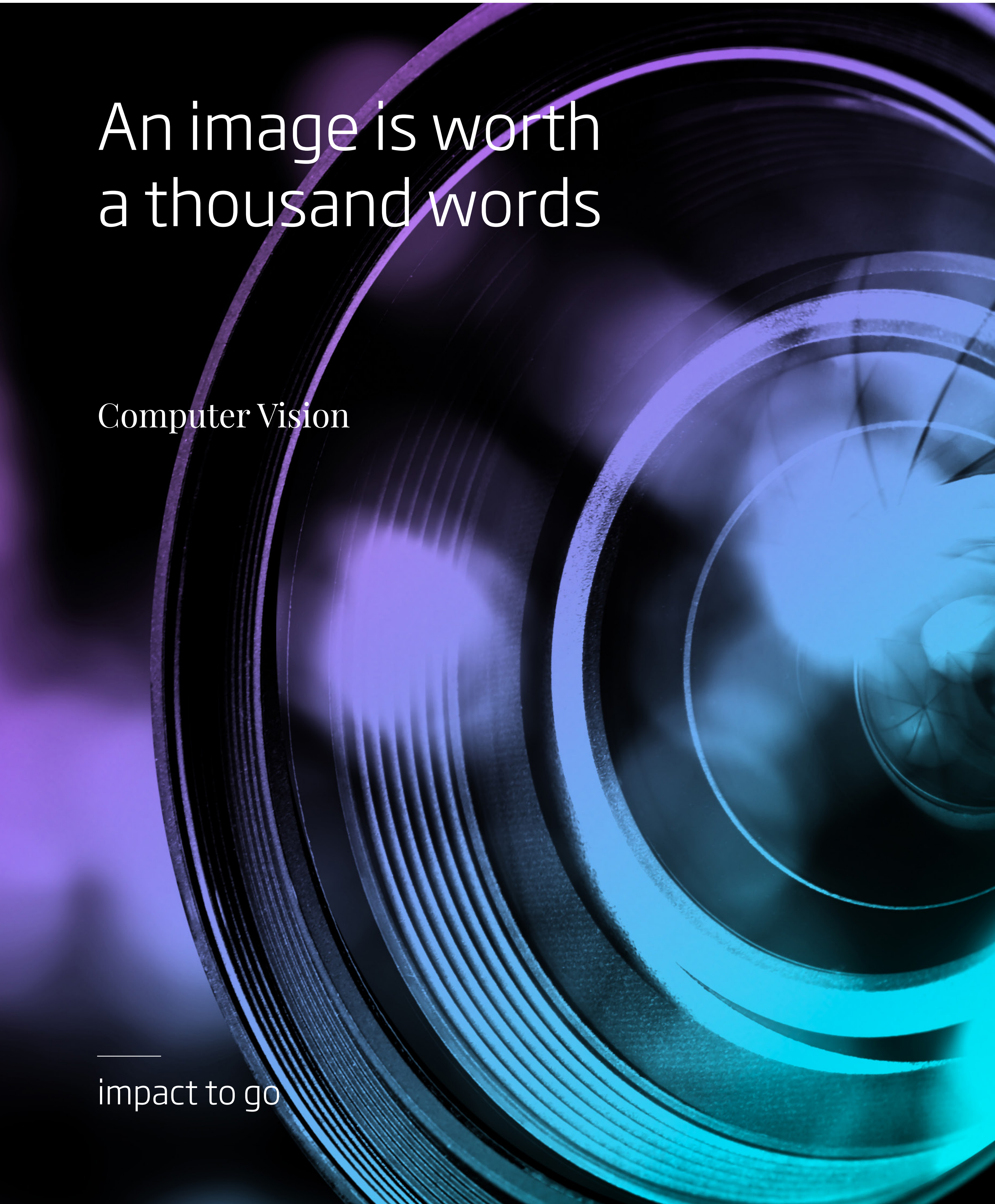


An image is worth a thousand words

Computer Vision

impact to go

Report presentation

“

In a world in which technology has become the main trigger of the new paradigm, one of the technological advances destined to generate the greatest disruption in the short-term is Artificial Intelligence—which has become a lever for processes automation and the development of new techniques. In particular, the rapid development of the Artificial Intelligence branch dedicated to interpreting images (Computer Vision) has produced disruptive solutions that bring important competitive advantages for companies that are opting for it in their digital transformation.

We at Minsait have the opportunity to be directly involved in this transformation, modelling the trends that will be decisive in Intelligence and Computer Vision, and working hand in hand with companies and organisations from all sectors to anticipate the new possibilities offered by this technology. Thus, we are leading the European R&D+i project, BeCamGreen, which is aimed at developing a Computer Vision and Artificial Intelligence solution to reduce city traffic and promote new sustainable mobility policies. We have also improved the efficiency of the production lines of various companies in the industrial sector, we have contributed to detecting fraud in the exploitation of natural resources, and we are also working to facilitate and speed-up public transport maintenance or the collection of waste thanks to Computer Vision, among others.

This report compiles the experience and opinions of different experts in Computer Vision and the related technologies and media (Big Data and Analytics, GIS, Satellite Images, Machine & Deep Learning) accrued from their participation in Artificial Intelligence and Computer Vision projects in numerous sectors of activity. With it, we want to share our views on: the key aspects and trends that companies must take into account whenever they plan, operate and develop their business using Computer Vision as a differential factor; the role of people in it; and, the strategies to be followed to successfully tackle projects that feature this disruptive technology.

Javier Torres Gutiérrez
Minsait

Executive summary

In the same way we use our eyesight to know what is happening around us, Computer Vision reproduces the same effect so that computers can understand the images they process and can act as programmed. While our eyesight is more adequate for qualitative interpretation of unstructured and complex scenarios, Computer Vision is more efficient in quantitative mediation of structured scenarios thanks to its velocity and precision.

To achieve accurate, reliable and repeatable results, Computer Vision tools include the intelligence needed to quickly and accurtttterns of objects of interest. There is an increasing number of applications in virtually all sectors of activity, among them:

Industrial processes

In a production line, for example, a computer vision system can inspect hundreds or even thousands of components per minute, making it possible to easily perceive object details that would otherwise go undetected by the naked eye.

Environment and energy

Computer Vision systems can be implemented in waste management plants to automatically detect the composition of waste thanks to specific cameras and vision software, and a robotic arm to classify it accordingly. It is also possible to create leakage monitoring systems in power plants to detect, for example, gas leaks invisible to the human eye in combined cycle plants.

Smart Cities

In the urban environment, it is possible to connect Computer Vision systems to existing cameras or to new ones implemented in public transport facilities to automate the collection of information and to prioritise incidents affecting urban furniture or infrastructures (e.g. road maintenance).

Computer Vision, through information analysis and processing, helps us better understand the environment that surrounds us.

Retail

One of the most important Computer Vision applications in this sector is the automatic control of warehouses and stock, using drones or other devices featuring cameras and computer vision software.

Banking and insurance:

Some Computer Vision applications in this sector can be the intelligent management of document flows from scanned images or automatic video monitoring systems.

The introduction of this new technology brings a new way of adding value to companies and institutions, and these listed above are just the first examples of what Computer Vision can achieve. However, to realise its full potential, the hardware and software must be combined in an optimal way, the image processing techniques must be used appropriately and Computer Vision must be combined with other technologies such as Big Data and GIS.

01. Context

Today's world is undergoing unprecedented transformation thanks to a common disruptor: technology. In a context of rapid evolution, the latest trends in technological development seek to imitate the processes of the human mind (the direct application of which is known today as Artificial Intelligence). In this sense, trying to automate the logic of human thought requires the ability to feed on both structured and unstructured data (for example, visual information) in the same way that a person sees and interprets an environment through his or her eyes.

In a strict sense, Computer Vision is therefore the branch of application of Artificial Intelligence that uses images from a capture device (for example, a camera) and then processes and analyses them to find defects, identify elements, contextualise environments and, ultimately, draw conclusions. The image obtained can be static (photo) or dynamic (video), analysing separately the static images that it is composed of (frames). These functionalities allow taking decisions based on visual info in a way that until now was exclusive to humans.

The first Computer Vision applications date back to the 1960s, with two notable events:

- The development of Frank Rosenblatt's perceptron, the first algorithm regarded as an artificial neuron with learning abilities implemented in a specific hardware. It was capable of distinguishing simple figures such as triangles and squares, based on trial and error [1].
- The first attempts at connecting cameras to computers with the aim of describing the images viewed.

Later on in the 1970s, the first Computer Vision commercial applications emerged, such as optical character recognition (OCR). However, it was not until the 80s and 90s that Computer Vision took its current revolutionary form with the arrival of the Internet (which gave access to more data and images), along with the rapid evolution of processors–and, more recently, of graphics cards.

Timeline of Computer Vision

1951
First commercial computer. UNIVAC I goes on sale, with a weight of 7,250 kg, capable of 1,000 calculations per second and 5,000 vacuum tubes.

1957
The perceptron is invented. Frank Roosevelt, American psychologist, creates the first Computer Vision algorithm capable of interpreting simple forms.

1959
The MIT Artificial Intelligence laboratory is founded.

1966
First attempts to connect cameras to computers. The aim is for the computer to describe the images captured.

1972
First Smart Robot with Computer Vision. Development of Shakey, the first robot capable of autonomous reasoning, combining primitive natural language solutions and Computer Vision .

1975
First digital camera . Introduction of CYCLOPS from Cromemco, the first fully digital camera on the market.

1976
Foundation of Apple Inc. and launch of its first computer.

1978
Creation of OCR programmes for the blind. Kurzual markets the first software capable of recognizing text and dictating it.

1980
Developments in multilayer neural networks.

1991
World Wide Web. The global Internet network is publicly launched.

1995
Face recognition. First software capable of identifying people through image analysis.

2005
Google Maps is launched. Satellite images are popularised.

2008
First commercial application of Augmented Reality. MINI creates a paper ad that when presented to a computer camera creates a 3D model of a vehicle.

2014
TESLA Autopilot launch. Autonomous cars are commercialised, using Computer Vision techniques.

2018
Amazon GO opens to the public. Opening of the first store without checkout, heavily dependent on Computer Vision technology.

(1) What is Computer Vision – Post 5: A Very Quick History
https://blippar.com/en/resources/blog/2016/10/04/what-computer-vision-post-5-very-quick-history/

Today, several factors have contributed to making Computer Vision undergo an exponential evolution, becoming an important part of the digital transformation of various sectors:

- The rapid evolution of hardware has brought us devices of greater computing capacity, better specifications, smaller size, weight and power consumption, and increased affordability, all of which has made it possible to implement new solutions that were impossible or not feasible until now. An example applicable to Computer Vision is the higher quality and resolution of cameras, including the ability to extract visible and non-visible light spectra.

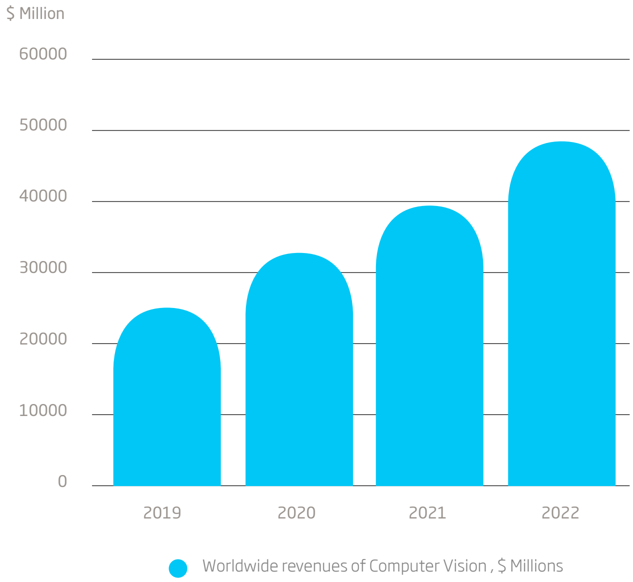
- As regards hardware, the sensors have also improved their specifications, such as resolution, accuracy and speed of response, in addition to increasing their control over environmental factors and their adaptability to them, such as the adjustment to the amount of light. However, the cost of some devices (such as laser-guided ones) has favoured the exploration of cheaper and more flexible alternatives, encouraging the development of Computer Vision techniques based on digital images.

- The improvement in processing capacity of devices has been reinforced by new calculation algorithms that have evolved through time, leading to more complex neural networks, such as Machine Learning and, in particular, Deep Learning, improving their capabilities and making them more efficient thanks to sophisticated programming techniques and data structures.

In fact, the figures show the rapid growth of this technology: it is expected for Computer Vision sector revenues to reach approximately $50 billion by 2022 (2).

One of the main reasons for the Computer Vision hype is its immense potential to automate processes and to generate new procedures that were not possible until now. On the one hand, it enables a degree of automation when applied to typically manual processes, which could only be performed by operators until now due to the need for visual analysis (which did not guarantee full or objective inspections). On the other hand, new disruptive techniques are created, such as the ongoing monitoring of environments, both industrial and natural, intelligent traffic systems, or qualitative controls of the entire production.

Worldwide revenues of Computer Vision



In both cases, the ultimate goal is to replace human vision with Computer Vision in applications where it was previously impossible, thanks to new and more effective algorithms, standardising their implementation. That way, the most basic processes could be fully taken over by machines, leaving it up to human beings to perform tasks requiring human supervision and greater complexity.

These systems not only allow defects to be detected more effectively, but are also capable of maintaining a constant optimisation of processes, allowing them to "learn" from errors and implement solutions to solve them. The main advantage of having a system-based optimisation process that replaces humans is that it can be done automatically, more accurately and in a more integrated way with the rest of the industrial machinery, as all data and corrections are digitised.

All that entails clear competitive advantages for companies that implement Computer Vision in their day-to-day activities. For example, in industrial factories, production is faster and more flexible, reducing downtimes and set-up times, offering a higher quality finished product.

(2) Computer Vision Hardware and Software Market to Reach $48.6 Billion by 2022
https://www.tractica.com/newsroom/press-releases/computer-vision-hardware-and-software-market-to-reach-48-6-billion-by-2022/

Another clear example where Computer Vision supposes enormous progress is the medical sector, allowing faster and more detailed diagnosis of illnesses and injuries. It also allows the application of more personalised and effective treatments in less time.

It can also improve patients' recovery prospects, since in many cases Computer Vision enables the early detection of ailments and the application of corrective treatments. An example of this is the pre-diagnosis of diabetic retinopathy which can lead to blindness, among other consequences, if not treated in time.

Despite its numerous business applications, Computer Vision is also available to the general public. Many applications of this technology are already integrated in the daily life of many people. A good example of this is the application in vehicles with some degree of autonomy (as is the case with the intelligent braking systems incorporated in some vehicles), or the augmented reality in many discovery games, or active photo filters, where Computer Vision is a fundamental part of the system.

However, and despite its rapid evolution, the challenge that Computer Vision faces before realising its maximum potential is to interpret not only images but also the environment and context in which they are found. That is the biggest challenge in the development of these systems: to ensure that they have the ability to interpret what is happening in an image (i.e. its situation) beyond finding and analysing points of interest such as defects or specific objects. Up to now, the ability to interpret was only exclusive to human beings. The future of Computer Vision is, therefore, to achieve the ability to contextualise images and, based on this, make the right decisions.

02. Challenges of Computer Vision

Technological challenges

- Evolution of hardware and software
- Amount of images needed
- Communication models
- Processors

Regulatory challenges

- Need for specific and adapted regulations
- Privacy of people in the images (analysis, filtering, pixelation, etc.)
- Regulated data storage

Organisational challenges

- Lack of knowledge/scepticism regarding the possibilities of this technology
- Possible organizational changes derived from the implementation of CV solutions

Today, Computer Vision has a huge transformation potential in the daily operations of various sectors. However, a set of challenges must be addressed before this technology can be fully exploited and commoditised:

Technological challenges

The current advanced stage of development of hardware and software has allowed to develop Computer Vision solutions that were previously unattainable, for example, achieving recognition under very adverse conditions (such as rain or darkness). These elements continue to evolve constantly and each development opens up a new range of possibilities applicable to this field.

In the field of software, one of the challenges is the amount of images and examples needed to "train" the systems. The advances in algorithm development are enabling performance improvements or more sophisticated learning in the case of neural networks, which requires less images and examples to train the systems.

In addition, communication models, which currently create barriers in the development of Computer Vision solutions (for example, the lantency or the insufficient bandwidth to transmit images of considerable size), are also constantly being improved.

As far as hardware is concerned, the improvement in the frequency and size of processors is decreasing, which in some cases forces processing to be done on servers rather than on the image capture devices, with the data flow that this entails. The still incipient quantum computation could be the solution to this problem.

Regulatory challenges

Artificial Intelligence requires a regulatory framework adapted to it, and although the first sketches are beginning to appear in this sense, there is still a long way to go. The specific case of Computer Vision needs images for its operation, and it is important to consider aspects such as people's privacy, an area in which the legislation in force still leaves room for improvement. For example, it is required to filter personal information in public environments (with techniques such as automatic detection and subsequent pixelation of faces), or warn whenever image recording is required.

The storage of images and personal data also plays an important role in regulatory terms. For example, the requirements of the Official Data Protection Act (LOPD), or the General Data Protection Regulation (GDPR) vary according to the type of data stored.

Organisational challenges

One of the hurdles for the successful implementation of Computer Vision is the lack of knowledge about the possibilities of this new disruptive technology that could provide solutions to traditional problems that occur in the day-to-day operations, or significantly improve their processes, by providing a differential value hardly comparable to other solutions. Understanding its potential is the first step to achieving its adoption. In the past, many companies found it difficult to implement these systems at a time when Computer Vision, due to insufficient hardware capacities and lack of optimised software, made it impossible to achieve the optimum results. All these difficulties have made companies become reluctant to implementing this tool again, despite the significant progress made in recent years. The possibilities offered by Computer Vision today are sufficient for companies to rethink their implementation in those processes where it was previously ruled out.

Computer Vision will improve the quality of those processes in which it is implemented. That may involve replacing the activities of some of the operators involved in these processes, allowing them to focus on more valuable and specialized tasks. The reluctance to adopt Computer Vision in this area is not justified: the total number of personnel devoted to a production process need not be affected, but the real change is in the increase in quality and efficiency of the tasks performed once Computer Vision is implemented. The personnel can be repositioned to handle tasks that were previously being neglected due to lack of resources, or to supervisory tasks (in processes where Computer Vision is not implemented or where there is no monitoring).

03. Computer Vision enablers

The introduction of this new technology brings a new way of adding value to companies and institutions. In this regard, there are several enablers to realise the full potential of Computer Vision:

- **Combinación adecuada de hardware y software:** the closed software used until now was programmed to be implemented in specific applications and with very specific hardware, which meant that Computer Vision implementation would operate under very limited conditions. The use of more open and flexible software brought about the adaptability to all types of hardware, in turn requiring less specialized and simpler devices (such as lower resolution cameras). That is why an optimum software-hardware combination can achieve greater flexibility when implementing Computer Vision solutions, also generating significant savings in equipment and devices, and making it possible to invest in improving the software.

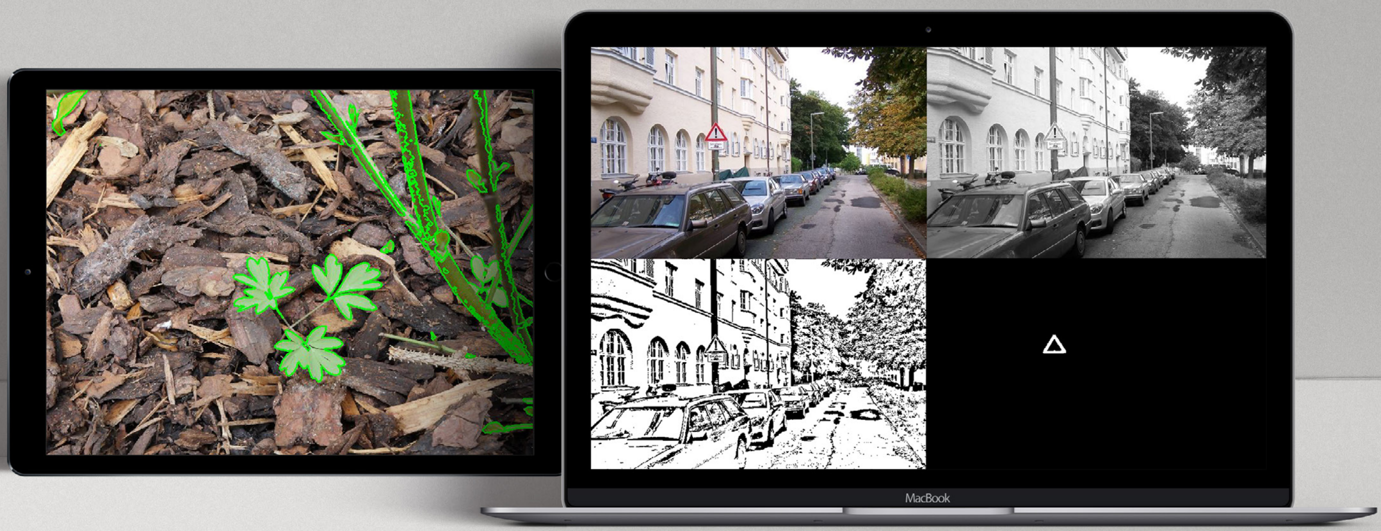
  Such is the case with drones, where LIDAR sensors are used to detect obstacles and distances by laser. However, this makes them heavy, with little autonomy and high cost, disadvantages that can be solved thanks to the use of Computer Vision, which would replace the sensors. The 2d images that the drone perceives can be processed, measuring in 3d by extrapolating the z coordinate, even if the captures are taken by lower quality cameras.

Real-time processing is especially key to sectors such as industry, where reduced production times imply reducing costs, making it possible to expand product lines, cover wider market areas and update products at the pace of technology development. Also the increase in speed enables the detection of production delays, with margins of up to milliseconds. To achieve these low execution times and high speeds, two enablers play a key role: image processing and segmentation techniques.

- **Image processing:** the use of correct image processing techniques, in which colour models, filters, brightness and shine are combined, are essential to providing solutions for the problems that may arise in the different use cases. Taking care of an image, modifying it and transforming it to achieve the expected results is of vital importance in these technologies, and can be a real differential value compared to other solutions with other techniques. For example, the incorporation of these techniques as a pre-process before Deep Learning solutions allows the data set required to train the algorithm to be much smaller, meaning savings in time and costs.

One of the key objectives of the Computer Vision implementation is to achieve real-time processing speed, which implies low execution times.

- **Segmentation:** it consists on defining in a concrete way each of the contents of the image, thus avoiding having to process all the content. This technique is another differential value that has high importance. With it, only the relevant parts are analysed, thus reducing the weight of the image, the necessary computing capacity and, consequently, the execution times.



I1. Example of segmentation, isolation of leaves within an image for further analysis (e.g. determination of plant health)
I2. Application of filters to detect a traffic signal

- **Combination with other technologies:** Computer Vision greatly expands its possibilities when combined with other technologies and connected adequately with the available technological means. That way it is possible to develop an End-to-End solution based on Computer Vision.

Enabling means

Deep Learning

- Incorporation of learning to systems
- Essential in non-deterministic processes

Capture devices

- Great variety of image capture devices available
- Application of already implemented devices / use of new hardware

Computer Vision

Synergies with technologies

GIS

- Incorporation of geographic coordinates to the data obtained visually

Big Data & Analytics

- Contribution of information from vision to the databases
- New and improved analysis models

End-to-End solution

Synergies

There are multiple technologies that a Computer Vision system can benefit from, and which can help develop a more complete and integrated solution that covers all requirements and provides the maximum functionalities possible. Among them, Big Data and GIS technologies stand out.

In the world of Big Data and Analytics, Computer Vision acts as one more data input, another lever in the analysis. It enables adding the information embedded in the images (not only figures) to the databases, thus providing new data sources that have not yet been exploited. The same data analysis model interprets and decides the weight of the data obtained from Computer Vision. An example of this combination are the models that detect unsold-products in retail companies. Rethinking them using Computer Vision can improve their performance, allowing them to know data and conclusions in real-time, instead of obtaining results once the sales persons have already detected the defect in sales.

A Geographic Information System (GIS) component can also be added. This technology integrates information systems in which data has a spatial component. In other words, it enables adding coordinates to data so that both the event and data are stored, along with the location. This generates more robust models, allowing, for example, analysis of population distribution in a given area. It also makes it possible to develop models with predictive capabilities. That is the case of monitoring critical masses: in large capacity buildings, such as stadiums, the behaviour of people entering and exiting events can be monitored, and the historic data enables us to predict the movement of masses during events. This can also be applied to traffic, predicting road conditions on certain dates based on the historical data analysed with Computer Vision.

Resources

Computer Vision makes use of different means for its operation, either to incorporate functionalities such as automatic learning (Deep Learning), or simply to obtain information (capture devices).

Applying Deep Learning to a Computer Vision system can teach a vision system to interpret changing and variable scenarios (i.e. non-deterministic) and to establish relationships between them. An example of this is the labelling of elements perceived by Computer Vision (e.g. such as knowing how to distinguish between the elements of an environment, what is a tree and what is not, taking into account the varied morphology of its species).

On the other hand, there is no need to apply Deep Learning in deterministic processes, where the set of options is very limited and the analysis context is easy to interpret. An example of this is the detection of defects in the industry, where all products are similar and where vision systems must focus on finding deviations from a perfect model. Due to the additional cost of programming and teaching such algorithms, it is best to use Deep Learning only when strictly needed.

Computer Vision is fed by images so, choosing the means of capturing them is a critical aspect. In this sense, there are many opportunities, which must be used on a case-by-case basis– from the use of existing camera systems, to later apply vision techniques; to the use of new technologies available, such as drones or satellite images.

For example, the commercial availability of high-resolution satellite images has made it possible to access a large amount of information, which can be automatically analysed using Computer Vision. That enables obtaining useful results such as possible patterns or the identification of elements from the air, which would otherwise have been hard to obtain using old techniques such as remote sensing and photo-interpretation. These techniques lack intelligence, and only have the capacity to identify simple parameters such as the determination of green areas by colour. The new Computer Vision capabilities make it possible to improve these processes, obtaining more accurate and reliable results (e.g. determining illegal border crossings automatically, analysing the history of satellite images of areas of interest). This input of intelligence is applicable to any capture device, hence the importance of choosing the one that best suits the circumstances.

04. Applications and use cases

The potential of Computer Vision is enormous, and the opportunities opened up can give companies the competitive edge in their sector. Although these systems can be implemented in any process requiring a visual inspection (which means a wide range of applications), there are sectors in which the convenience, opportunities, and advantages of Computer Vision stand out over others. These sectors are described below:

Industrial processes

Visual quality control
Detection of deformations, cracks and/or breaks in industrial parts

State of Welds

Paint quality

Production process control
Control of traceability of parts in transport (transit)

Hyperspectral control
Quality control of fresh food

Environment and Energy

Waste management
Waste classification for recycling

Detection of waste on site

Fraud
Control of fraud in subsidies

Illegal mining and fishing

Illegal constructions

Water Consumption Control

Farming
Plant health and pests

Ground conditions

Agricultural machinery Automation

Pruning Alerts

Installations
Heat and gas leaks in installations

Natural disasters
Fire detection

Monitoring
Inspection of areas difficult to control

Automatic control in marine rescue

Smart Cities

Urban maintenance
Faults in public roads

Transportation infrastructure control

Mobility
Security control in critical areas

Traffic and parking situation

Vehicle occupancy control

Autonomous cars

Urban town planning
Intelligent 3D paps

Energy efficiency in buildings

Retail

Stock control / logistics
Warehouse status control

Retail areas
Customers flow control and trends

Incidents surveying

Banking and Insurance

Documentation management
Document classification

Loans
Personalised financing

Insurance
Automatic expert reporting

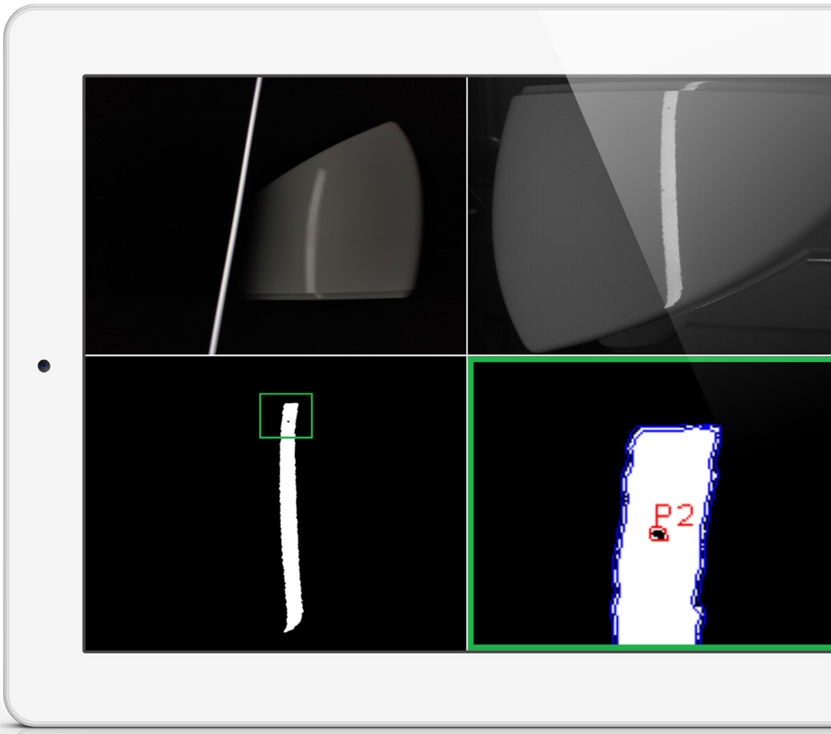
4.1 Industrial processes

As opposed to other sectors where the use of Computer Vision is still incipient, the industrial sector frequently uses the technology as a lever to optimise production processes, which are very demanding as regards time and costs per unit.

The main objective of the application of Computer Vision in the industry is, therefore, the automation of processes, and any type of industry admits to Computer Vision automation to a greater or lesser extent. In addition to improving productivity and performance, Computer Vision can also contribute to the monitoring of operators and machines. It can detect an error so that actions can be taken accordingly, for example, removing the defective unit and determining where the fault has occurred, controlling that all movements are correct with the ultimate goal of reducing time and costs, and performing more efficiently.

Apart from the quantitative aspects, Computer Vision also contributes to improving the quality of production processes. It is capable of detecting errors with greater precision and speed than humans, which translates into a better quality service, and a greater satisfaction of the final customer, who will receive a defect-free product.

Within the sector, Computer Vision can be implemented in several areas:



Visual quality control is not currently automated in most cases. This means that solutions such as sampling are used in order to keep pace with the rest of the production, which means that reliability is lost because not all the parts produced are analysed. However, the growing demand for product quality, as well as in production processes, renders these solutions obsolete. Computer Vision enables inspection of 100% of the products.

The examples in this sense vary from the inspection of the state of the nuts–detecting errors such as deformations, cracks or breakages using image-processing techniques and algorithms, and which has been in place for some time–to other more recent examples such as the classification of tomatoes, up to now carried out by sensors and which can now make use of Computer Vision to separate by colour.

Another of the most frequent use cases is in quality control of welding processes, making it possible to detect defects such as porosities, openings or transfers in the weld bead. Defects in paints are also detected by Computer Vision, obtaining greater control over inspection of parts by passing a light beam and subsequently analysing the image, friction, dirt or imperfections found.

Comparison of costs in inspection of paints in the factory

	Without Computer Vision		With Computer Vision	
FTEs required for inspection	12	-60%	5	
Turn queuing displays	3		3	
Annual cost by FTE and shift	€45.000		€45.000	
Total annual inspection costs	€1.620.000	€-945.000/yr	€675.000	
Quality-related sanctions per year	6		0	
Average cost per claim	€3500		€3500	
Total annual cost for penalties:	€21.000	€-21.000 €/yr	€0	
% scrap material	15%	-2%	13%	€1560 M/yr in prod.
Annual total cost of scrap	€1.385.000	€-185.000€/yr	€1.200.000	

Total savings in production due to implementation
€1.151.000

—

Implementation costs
Of Artificial Vision
€212.500

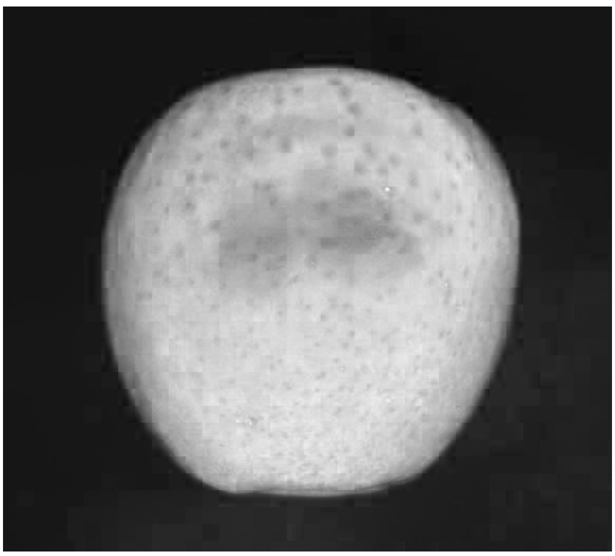
=

First Year Total Savings
€938.500

Total annual savings
(from the 2nd year)
€1.151.000

Computer Vision can also be used in the **control of the production process**, checking the production whenever a failure is detected, in order to detect and solve the origin of this error. The application of traceability to the parts can help find the points of interest and create a unique identifier for each part. That, in turn, generates automatic quality control processes for the industrial flows, providing visibility of defective parts in the production line thanks to advanced image processing. Such traceability enables the user to know what machine is causing the defects or even if the problem is outside the factory, for example, during transportation. This, in turn, means an ongoing improvement of the production process through the application of corrective measures.

Beyond the visible defects, Computer Vision also makes it possible to identify patterns that are invisible to the human eye. Thus, the use of **advanced hardware (such as hyperspectral)** allows, for example, to detect factors that can determine the quality of fresh foods such as eggs or meat and fish, finding chemical contaminants and detecting possible bacteria that have characteristic spectral signatures. In addition, the contactless inspection of food reduces the possibility of contamination, thus achieving a higher quality process.



4.2 Environment and energy

The Computer Vision applications in the natural environment are linked to aerial images of large areas obtained by satellites and of more limited areas captured by drones, as well as static images of a given environment taken by fixed cameras.

In the case of satellite images, there has been significant evolution in recent years, following the commercial availability of high-resolution satellite imagery. As a result, a vast quantity of information is now accessible. This volume of information is too large for a human supervisor to analyse it in its entirety, but can be processed and analysed by a Computer Vision system. On a smaller scale, the popularisation of commercial drones has extended the application of these systems in more limited environments. The problem of high costs and low horizontal detail of high-resolution satellite images has been overcome, which previously prevented the viability of many use cases.

Along with the increased availability of images, the recent progress in hardware has brought great advances as regards environmental issues.

For example, specific hardware such as hyperspectral can detect anomalies that would otherwise be invisible to the human eye (such as gas leaks, or automatic recognition of material composition).

Therefore, Computer Vision is applicable to a wide spectrum of activities related to the environmental sector:

In waste management, on the one hand, Computer Vision can be used for the classification of waste in recycling, automating the waste sorting process by type of material, separating it with robotic arms. This replaces the current methodology, with magnets and sorters, which makes the process slow and inefficient, and significantly improves the recycling rate. On the other hand, waste can also be detected in the environment in the form of pollution, gases, foams, marine debris, or illegal waste dumps.

Computer Vision is also applied for **detecting environmental fraud**. For example, it is possible to detect false olive tree plantations created to collect European subsidies for the volume of olive trees planted.

Computer Vision, drones and satellite images allow us to detect illegal constructions or expansions or undeclared pools, and with the help of GIS it is possible to locate the event geographically. Water consumption fraud is also detectable by comparing data on the filling status of private pools or the colour of the grass of the gardens to theoretical consumption data. These applications optimise agent on-site inspections and contribute to achieving 100% effectiveness in fraud detection, with more precise and robust models.

These technologies also make fraud detection possible at a larger scale, which might otherwise go undetected: an example of this is the detection of illegal fishing on large areas of water, or of illegal exploitation of natural areas of land, such as unauthorized mines.

Agriculture is a sector where Computer Vision has many applications. It can be applied to the monitoring of plantations, detecting plagues, controlling very important characteristics of soil during cultivation (such as humidity and fertility), counting plantations and controlling the perimeter.

In some cases, agricultural harvesting is done by means of automated machinery with many sensors, which requires very high investment. These sensors can be replaced by Computer Vision which, by means of image processing using colouring, and if necessary by adding learning, could perform the same function. In addition, the condition of the trees can be automatically monitored to detect when pruning is necessary.

Thanks to the non-visible spectrum cameras, any gas, water or heat leaks in **installations** (such as leaks in combined cycle plants) that are not visible to the human eye can be detected and located with the help of GIS. This type of solution enables users to schedule maintenance and detect, for example, heat leaks in power lines by drones, avoiding the traditional and expensive inspection flights in light aircraft.

Anomalies and natural disasters can be monitored through continuous observation of different spaces, and alerts can be configured in case of an unforeseen event.

The surveillance of large natural parks, or even fire control, could be entrusted to Computer Vision. It would be able to detect the initial smoke and warn the nearby population that may be in danger as well as the fire fighting teams.

In addition, with this technology, and by means of capture solutions, it is possible to automatically monitor areas that are difficult to inspect, or that present dangers and risks for people, and thus guarantee the safety of the personnel involved. The latter has more fields of application, such as safety in the Industry, or in sea rescue.

4.3 Smart Cities

A Smart City is a connected city, where all the services provided by the city council to citizens to improve their quality of life are managed efficiently and transversally, based on the multi-connectivity of infrastructures and using the capabilities offered by new technologies. These services are adapted and personalised based on the reality of citizens, so Computer Vision has great application in this field:

In urban maintenance and inventory keeping, problems in urban installations and furniture can be identified, including the need for maintenance in public parks, streetlamp breakdowns, poor condition of containers or holes in the public thoroughfare. For example, a series of cameras located on buses can detect and classify the problem; and GIS technologies, by linking geographical location, can detect incidents automatically, prioritising them and avoiding the need for the citizen to report via smartphone. In this way, the efficiency of the system is improved.

The maintenance of transport infrastructures, such as railway tracks and cables can be programmed automatically by implementing Computer Vision in the trains themselves. Other monitoring possibilities, such as the maintenance of radar huts at the appropriate temperature, can be solved thanks to this technology, controlling the impact of the sun on the huts.

In behaviour monitoring and incident prevention, Computer Vision complements the current security systems, making an efficient use of cameras already installed and applying intelligence to their images and videos to detect unforeseen events and generate alerts in real-time and automatically.

This is especially useful for emergency centres to anticipate any situation and automate the generation and sending of alerts.

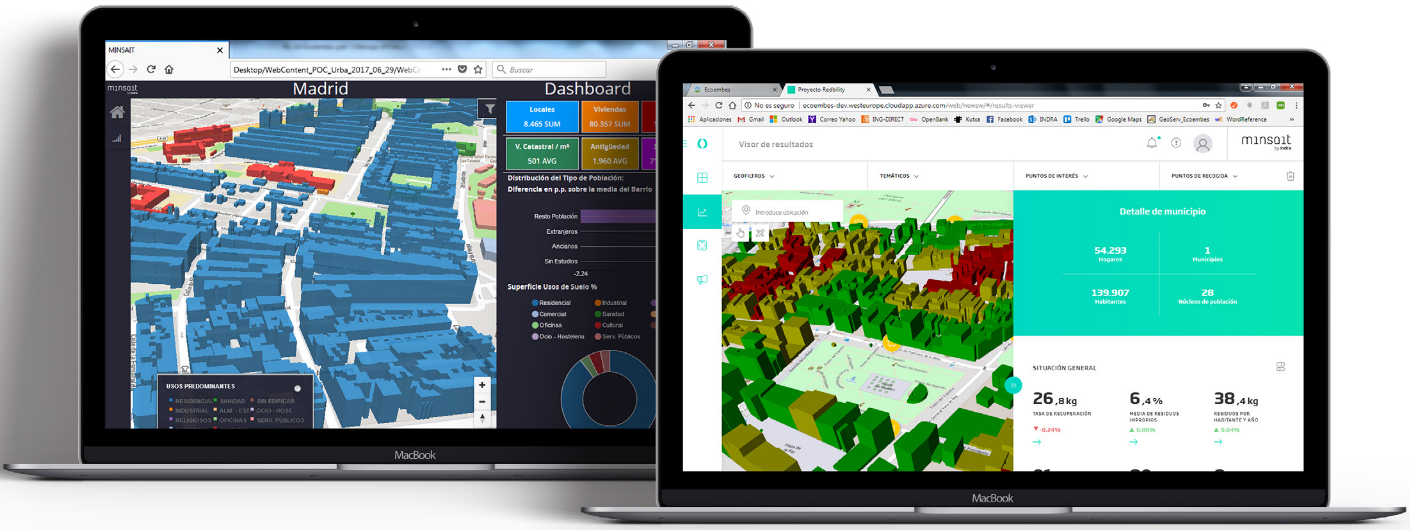
In this field, there are control applications for crossings and level crossings that generate warnings in case of obstructions, thus preventing accidents.

In mobility, the data obtained can be used so that citizens can have real time information of road traffic or the availability of parking before using the car. That way, it is possible to avoid traffic jams, crowded areas or even air pollution spikes.
The reduction of accidents by means of autonomous cars is also a reality thanks to Computer Vision, since the guidance systems that these vehicles have are based on this technology: for example, all lane change warning systems, or cruise control variable speed depending on the preceding vehicle are Computer Vision applications.

Computer Vision also complements the current security systems, making it possible to take advantage of the surveillance cameras already installed and applying intelligence to their images and videos. That way the system can detect unforeseen events and generate alerts automatically and in real-time. This is especially useful in emergency centres, anticipating to any situation and automating the generation and sending of warnings.

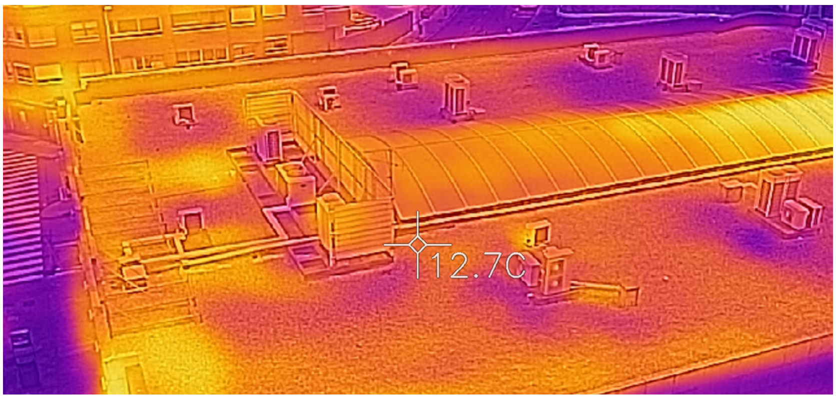
In urban planning, the traditional tendency for large cities to expand has been replaced for urban modification and modernisation. All of these changes, such as the creation of parks, building remodelling or road diversion can be previewed and monitored beforehand thanks to Computer Vision. To do that, Smart 3D maps are created in which all the elements are labelled or grouped into categories (trees, buildings, benches, etc.), allowing for element segmentation and an automatic updating of plans when changes are made to a city layout.

This is the main contribution of Computer Vision: up until now, non-intelligent 3D model generation technologies, such as LIDAR, have been used, without classifications or recognitions of any kind.



In addition, thermographic cameras can be used to monitor the energy efficiency of infrastructures and buildings, determining the heat they give off both during the day and at night, and identifying the areas with the greatest losses for subsequent repair. The terrestrial cameras, installed both in street furniture and in vehicles and buses, make it possible to check plans such as façades or window panes.

Another example of the application of Computer Vision in cities is in the field of Real Estate development, where it is possible to recognise house or building plans, generating a 3D model from them by means of augmented reality. The models generated are interactive, and allow the user to preview the possible structural or visual changes that need to be made in the buildings and floors.



4.4 Retail

The popularisation of devices such as smartphones with cameras or, more recently, mobile drones and robots, as well as the latest developments in vision hardware–especially in cameras (which are already frequently used in industrial and commercial areas)–are opportunities to apply Computer Vision in the retail sector. It can be directly applied to enable an ongoing continuous and automatic monitoring. With solutions of this type, the retail company can improve its operations, and can retain greater control over its products and its customers' behaviour towards them.

Companies can obtain improvements in two main areas:

In stock control and logistics, the creation of models for existing products through the capture and processing of images enables the user to manage stock location and availability. A drone can be configured to fly indoors in the warehouse, create a 3D map of it and monitor it, monitoring stocks, categorising the inventory to suggest more efficient redistributions and generating warnings in case of anomaly. Thanks to the map generated, the drone learns from the environment and creates routes, avoiding obstacles without the need for additional devices such as LIDAR. All this makes an ongoing and real-time control of the warehouse possible in a fast and automatic way. The need for inspection staff to frequently analyse the space is thus eliminated.

In the monitoring of retail areas, Computer Vision and indoor location techniques can locate customers, establish flows, transits and hot areas to identify purchase trends, such as preferences and behaviour in relation to certain products that customers show interest in. The information gathered can generate heat maps and predefined routes for customers, allowing, for example, the strategic placement of products.

It is also possible to use Computer Vision solutions to resolve incidents in the commercial stores themselves. That way, automatic alerts are produced if incidents such as stockouts, queues or bottlenecks at checkouts occur, thus improving the consumer experience and business efficiency. Suppliers can also benefit from these solutions, enabling them to monitor the prices and shelf locations of their products relatively frequently. It is not always easy for suppliers to know the retail price of their products or their positioning as agreed with the retailers, so they are now able to avoid the costs of having to constantly visit supermarkets to supervise.

The application of Computer Vision in retail not only generates benefits for the business in terms of efficiency and control. Customers also perceive advantages of the implementation of these systems, which improves the shopping experience. The combination of Computer Vision and data analytics makes it possible to analyse trends, particular tastes and previous purchases, as well as the electronic devices used. It is possible to detect areas of special interest by monitoring the footfall, and by obtaining the identity of the customers and identifying their consumption habits we can make suggestions to them.

4.5 Banking and insurance

The insurance and financial services sectors can also benefit from Computer Vision technology. Some of the benefits include the automation of bank operations– such as the automatic capture and processing of relevant data, optimising internal processes and allocating human resources to more complex analysis tasks requiring qualified decision-making (instead of using these resources for repetitive and low-value tasks).

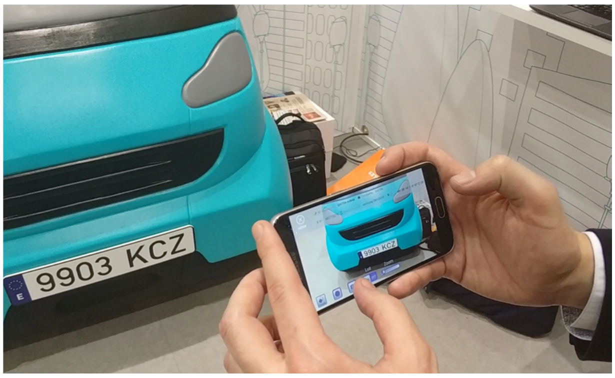
The most important use cases in the banking and insurance sector are:

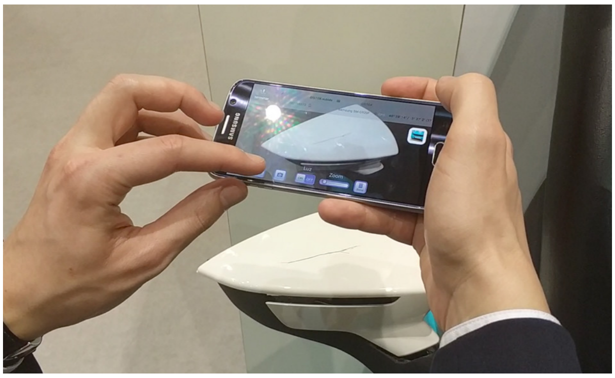
Automatic processing and classification of documentation: Computer Vision can use images obtained by automatic scanning to automate document workflows, classify and separate documents automatically and facilitate the extraction, validation and uploading of data to operating systems.

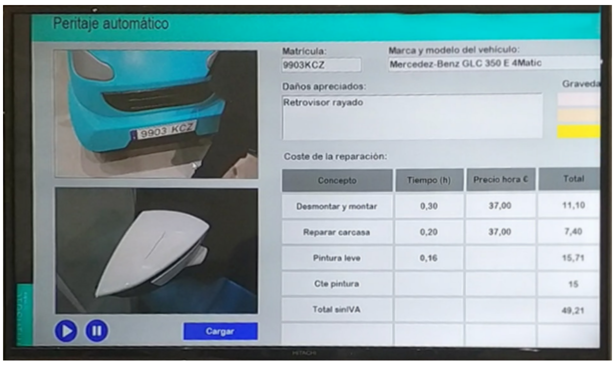
Personalised customer financing: uploading the photo of a vehicle to a mobile application enables the bank to identify the make and model, as well as its technical characteristics and price. With the information and the financial data of the customers, the bank can offer financial products adapted to the situation of each customer.

Expert reporting of assets: Computer Vision enables the customer to send information of any damage to his assets caused by accidents or factory failures, thanks to the camera of his Smartphone. The expert receives this information and, thanks to image analysis techniques, estimates repair costs in minutes, identifying the damaged parts and accessing a catalogue of parts through Cloud technologies. This type of reporting is used in areas like the automotive industry, where insurance experts can assess damage to vehicles without the need to pay visits.

It also enables innovative solutions to be implemented in the insurance area, like the almost immediate delivery of experts' reports using the Smartphone cameras.







05. Keys to successfully approach a Computer Vision project

In order to take full advantage of Computer Vision in each of its innumerable applications and use cases with a differential approach, and to add value to the company, three key aspects must be considered when approaching a project with this technology:

- Approach the project as an innovation project: It is important to bear in mind that Computer Vision is a disruptive technology, so its implementation should not be approached as a traditional IT project. Therefore, in the context of developing Computer Vision projects, it is worth considering new work methodologies–for example, Agile methodologies, in order to obtain viable minimum validated products and according to business requirements in a more controlled way–or alternative financial sources as innovation projects instead of IT projects as usual.

- Develop Integral HW-SW solutions: coming up with a solution that adds value to the company means developing it in its entirety, end-to-end. Software and hardware must go hand in hand, and the complementary technologies that are beneficial to the solution must be integrated at an initial stage. The development of a partial solution using a team of experts specialising in a single area does not realise the full potential of Computer Vision technology.

- Follow a clear, iterative and scalable methodology: a Computer Vision implementation project requires a phased methodology, including a first phase that analyses the current status and problems, a second phase for creating a multidisciplinary team, an iterative trial and error period to find an optimal solution, and a final rollout and scaling phase:

Computer Vision Project methodology

Analysis

- Diagnosis of the current situation, analysis of problems.
- Identification of opportunities for improvement and possible use cases with Computer Vision.
- Development and validation of a first Business Case.

Team creation

- Selection of the members of the Computer Vision project team, based on a series of recommendations:
  1. Members of different areas (multidisciplinary team) to achieve an end-to-end solution, with integration of other technologies and business vision.
  2. Integration of client staff to gain deeper insight into the current operations.

Iterative process

- First development of a pilot test of the solution, regardless of the operation of the company, to validate its suitability.
- Running of tests and adjustments iteratively to achieve an optimal solution.
- Gradual integration of the solution in business operations.

Scaling of the solution

- Full integration of the solution into the company's procedures.
- Scaling of the solution to other sites/other areas of application.

Strategic follow-up through a **Project office**
Strategic considerations on Computer Vision to get value out of their application
Monitoring with the direction of efficiencies objectives and of lines of work

06. Conclusions

Computer Vision has evolved exponentially in recent years, thanks to advances in the devices, processing speeds and software. The rapid development has positioned this technology as a tool within the current digital transformation of companies, offering disruptive and value-adding solutions in a wide variety of sectors and particular use cases.

07. Acknowledgements and authors

Authors:

Javier Torres Gutiérrez

Senior Manager Computer Vision Practice
jtorresg@minsait.com

Rafael Rivera Retamar

Senior Consultant Computer Vision Practice
rrretamar@minsait.com

Javier José Valero Gómez

Senior Consultant Computer Vision Practice
jjvalero@minsait.com

Collaborators:

César De Andrés López
cdeandres@minsait.com

Sergio De Mingo Castillo
sdemingo@minsait.com

José Carlos Rodríguez Mora
jcrmora@minsait.com

Marco Einöder Moreno
meinoder@indra.es

Alejandra Cristina García Hooghuis
acgarcia@minsait.com

Damián Andrea Santoro Martínez
dasantoro@minsait.com

Pedro Amalio Orihuela Gómez
paorihuela@minsait.com

minsait
by Indra

impact to go