



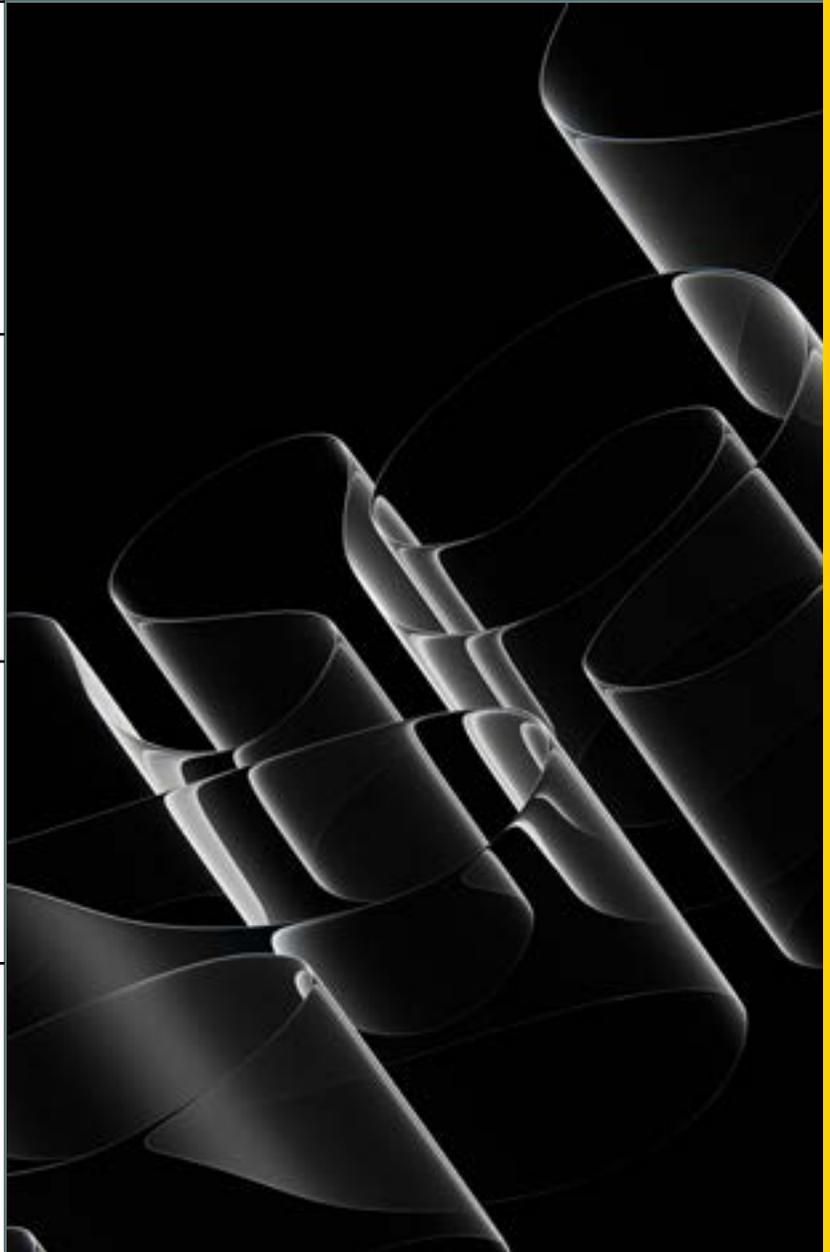
Highlights

**Computer Vision
engineering:
what machines
can see**



**PROSEGUR
SECURITY**

Index



01

Seeing to understand: The origin and concept of computer vision



Computer vision is the **ability of a machine to interpret and understand the visual world based on different algorithmic models to extract meaning**. Although it may sound like cutting-edge technology today, its roots date back to the early decades of modern computing. In the 1960s, researchers at MIT were already dreaming of equipping machines with digital eyes capable of **“seeing”** their surroundings. However, it was not until well into the 2010s, with the deep learning revolution, that this discipline began to be deployed on a massive scale, making its way into real-world applications.



→ But what does “seeing” mean for a machine?: The machine vision

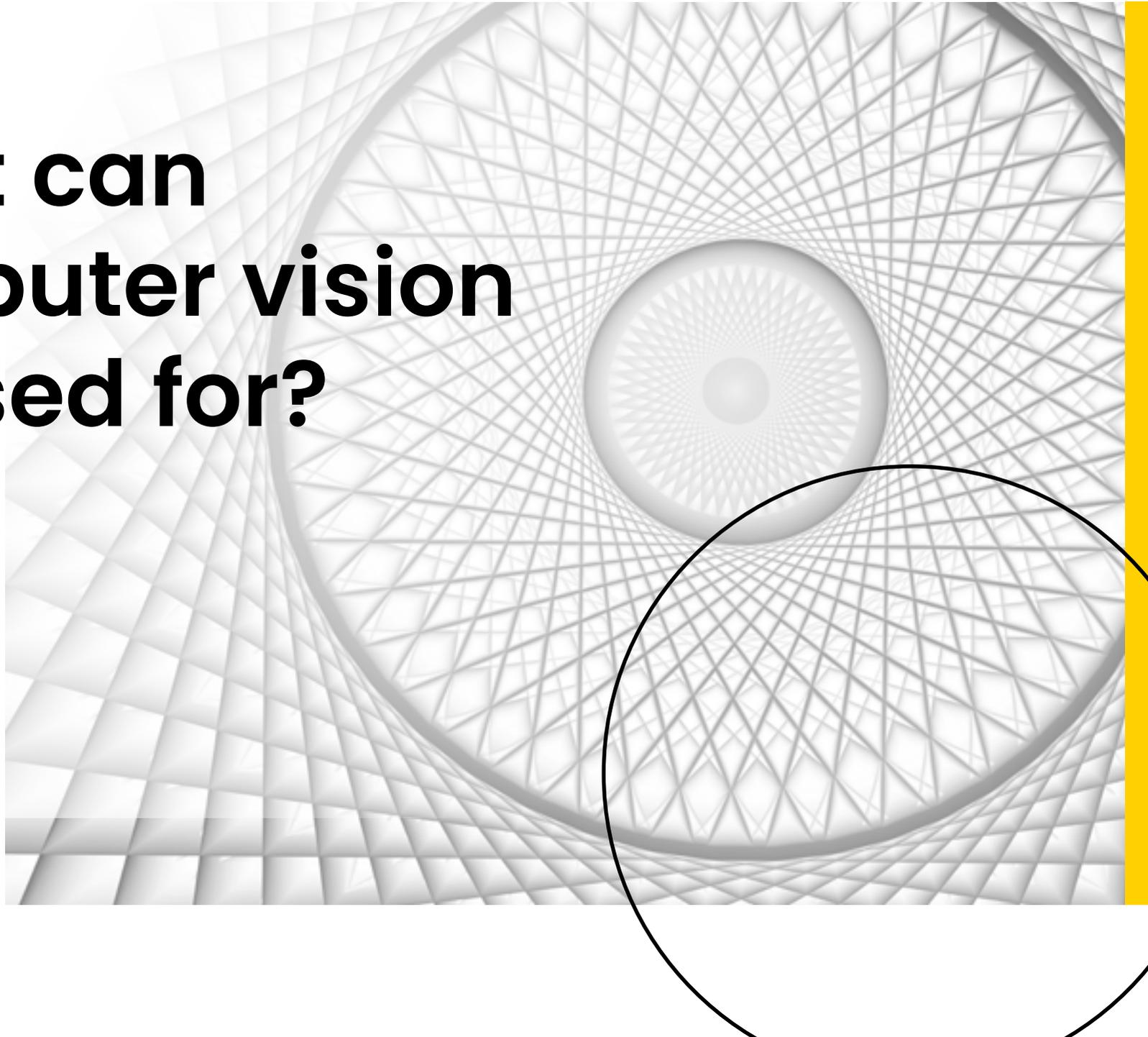
Essentially, this technology aims to convert images —sets of pixels with no inherent meaning— into understandable and useful information. To achieve this, computer vision relies on a processing chain that mimics, in its own way, the human visual system, although it operates through radically different processes: it first captures images through sensors (such as cameras), then analyzes them to extract visual patterns (edges, shapes, textures), and finally interprets them by assigning meaning based on context.

This process combines optics, mathematics, statistics, artificial intelligence, and neuroscience. It is not enough to recognize an object;

the system must understand what it is, what it is doing, and how it relates to its environment. That is why computer vision is not just about “seeing,” but also about **processing and interpreting what is being detected**. This distinction is key to assessing its transformative potential: we are not dealing with a simple evolution of security cameras, but the emergence of systems capable of **supporting and executing decisions in real-time** based on visual stimuli, anticipating risks, and operating autonomously in complex environments.

02

What can computer vision be used for?



Computer vision is a **versatile technology** with a surprisingly wide range of practical applications, whose influence extends across different sectors of society. Its potential lies not only in the capacity of automatize visual processes, but also in **significantly improving the accuracy and quality of decisions** based on the intelligent interpretation of images and videos¹.

One of the most immediate and impactful fields is **security**, where computer vision offers a **decisive strategic advantage by anticipating and detecting threats** with greater rapidity and accuracy than human capabilities. Through **behavioral pattern** analysis systems, this technology enables **real-time identification of potentially dangerous or suspicious situations**, facilitating agile and effective responses. Thus, people, data, and technology merge into an efficient and effective protection ecosystem in terms of time and resources.

¹ For further insight, we recommend "[Vision](#)" by David Marr.



Moreover, enhanced threat detection in monitoring environments is a notable aspect of the agility of algorithms, specifically in the detection of suspicious movements, intrusions, or anomalies. This makes computer vision a valuable tool that complements the essential work of security personnel, adding value to their efforts. The synergy between these technologies, which allow for the use of combined data with relevant human experience translates into a level of security never achieved, which we refer to as **hybrid security** at Prosegur.

A clear example of this would be the usefulness at a large-scale event, such as an international music festival or a sports competition, where algorithms allow us to guarantee the safety of attendees, through computer vision. Thus, artificial intelligence and computer vision algorithms allows to ensure safety of attendees by detecting weapons or suspicious behaviors earlier and better based on real-time information, which is processed at that very moment by the professional in charge of the camera, who can immediately inform the on-site security personnel.

That is why we say it empowers and with an increasingly high level of sophistication: for instance, we can now detect smoke through images much earlier than traditional sensors under suitable visibility conditions.

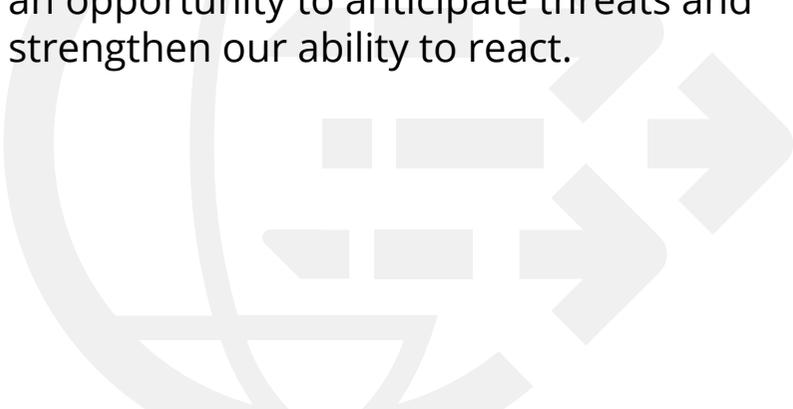
Other sectors benefiting from this technological advance include the **healthcare sector**, where computer vision improves diagnostic accuracy and speeds up disease detection through the automated analysis of medical images (e.g., X-rays, MRIs). These systems support **early and personalized diagnoses**, multiplying the effectiveness of treatments and helping to **save lives**.

In the field of **transportation**, real-time environment recognition propels the development of **autonomous vehicles** capable of identifying pedestrians, obstacles, or unexpected road changes, thereby **improving road safety and significantly reducing accidents in this type of vehicles**.

In industry, **automatic visual inspection** enables instant detection of **quality issues** or risks in critical production processes. This visual automation not only **reduces costs** but ensures **consistent quality standards**.

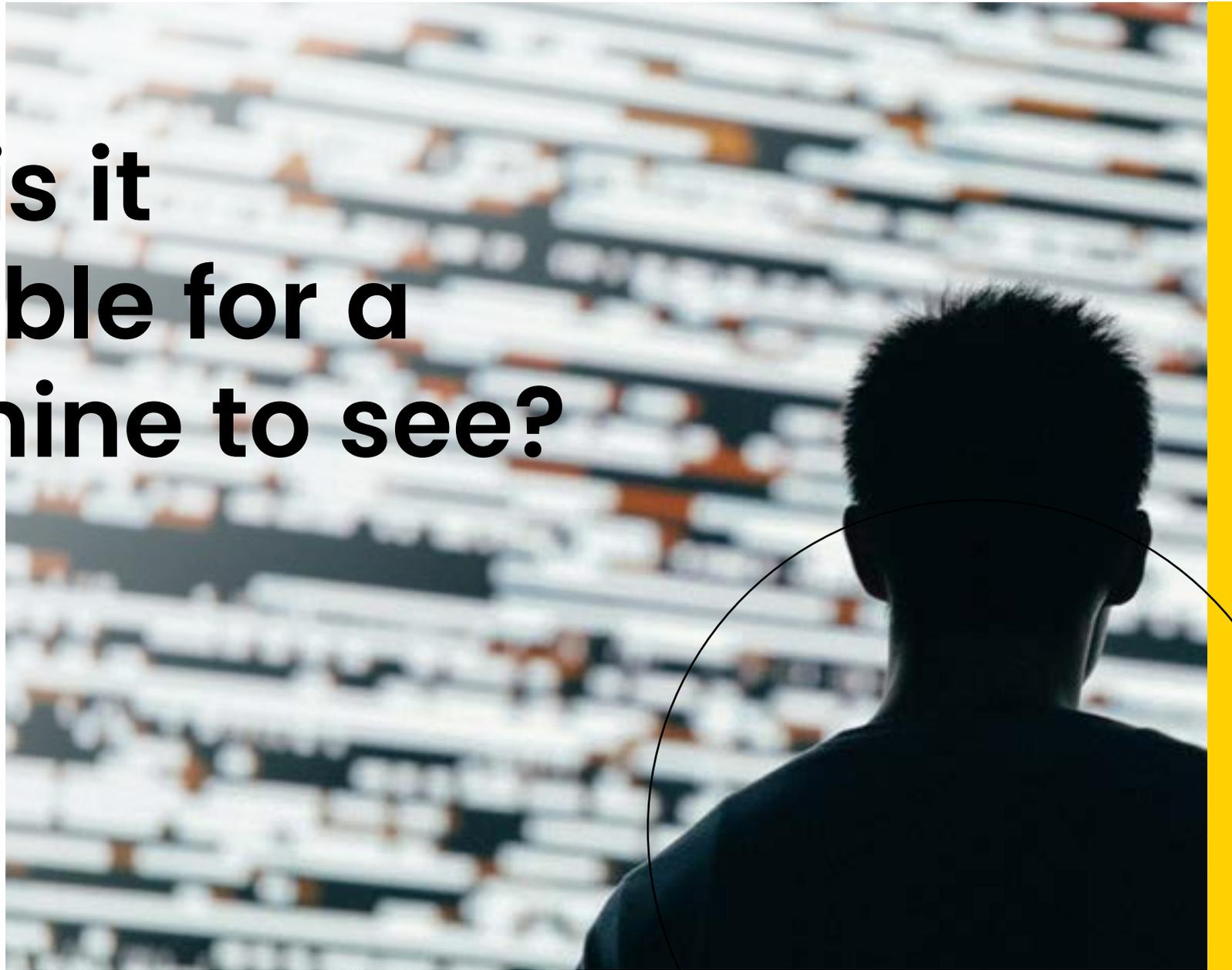
Finally, in **agriculture**, through the use of, for example, multispectral cameras, computer vision is used to monitor crops with pinpoint accuracy, identifying needs for water, fertilization, or pest control. Among its most widespread applications are plant identification and, in conjunction with robotic systems, harvest automation. In short, this technology promotes more efficient use of natural resources, boosting capabilities to achieve **sustainable and safer production**.

In summary, computer vision rises above a mere technological advance, where the synthesis of human ingenuity and algorithmic power creates an intelligent fabric capable of deciphering the visual world with superhuman nuance. The **alliance between the digital eye and analytical** intuition allows us to read the pulse of reality before the shadow of risk materializes, charting paths of protection that vibrate with precision and resilience. In this intersection of light and data, machines no longer just see, but also emulate human abilities of understanding and intuition, transforming every frame into an opportunity to anticipate threats and strengthen our ability to react.



03

**How is it
possible for a
machine to see?**



Computer (or artificial) vision greatly benefits from the ability of **Convolutional Neural Networks (CNNs)** to extract complex visual features, and more recently, from the versatility of Transformer-based models (such as Vision Transformers and multimodal **Vision-Language Models – VLMs**) to process sequences and relate visual information to the semantics of natural language. **CNNs are artificial intelligence models specifically designed to work with images: they function by applying a series of “filters”** that capture

local patterns (edges, textures) and, as they pass through successive layers, they construct increasingly complex representations (shapes and objects). LLMs, on the other hand, are systems trained on massive volumes of text —and, in their multimodal variant, with visual data as well— **capable of understanding, relating, and generating natural language, providing context and meaning to CNN outputs.**

Capture and preprocessing

The images are obtained from sensors (cameras, drones, satellites) and normalized to correct brightness, contrast and remove noise, ensuring homogeneous inputs for subsequent analysis.

Features extraction

CNNs deploy trained filters across successive layers. The initial layers highlight edges and textures; intermediate layers combine these elements into simple shapes; and the deepest layers assemble these shapes to recognize complete objects (e.g., a face or a vehicle). Each filter adjusts its parameters by learning from large datasets of labeled images, thus optimizing detection performance.

Contextual reasoning

Multimodal LLMs take the feature vectors generated by the CNN and link them to a vast corpus of text and visual knowledge corpus. This enables practical meaning to be assigned (e.g., identifying a “cargo container” instead of a “shaded rectangle”), external data (geographic location, historical records) to be integrated to enrich the interpretation and ultimately deliver outputs in natural language or in operational rule-based formats.

Output and action

The result combines precise metrics with semantic descriptions, enabling applications ranging from real-time security alerts to detailed instructions for autonomous systems.

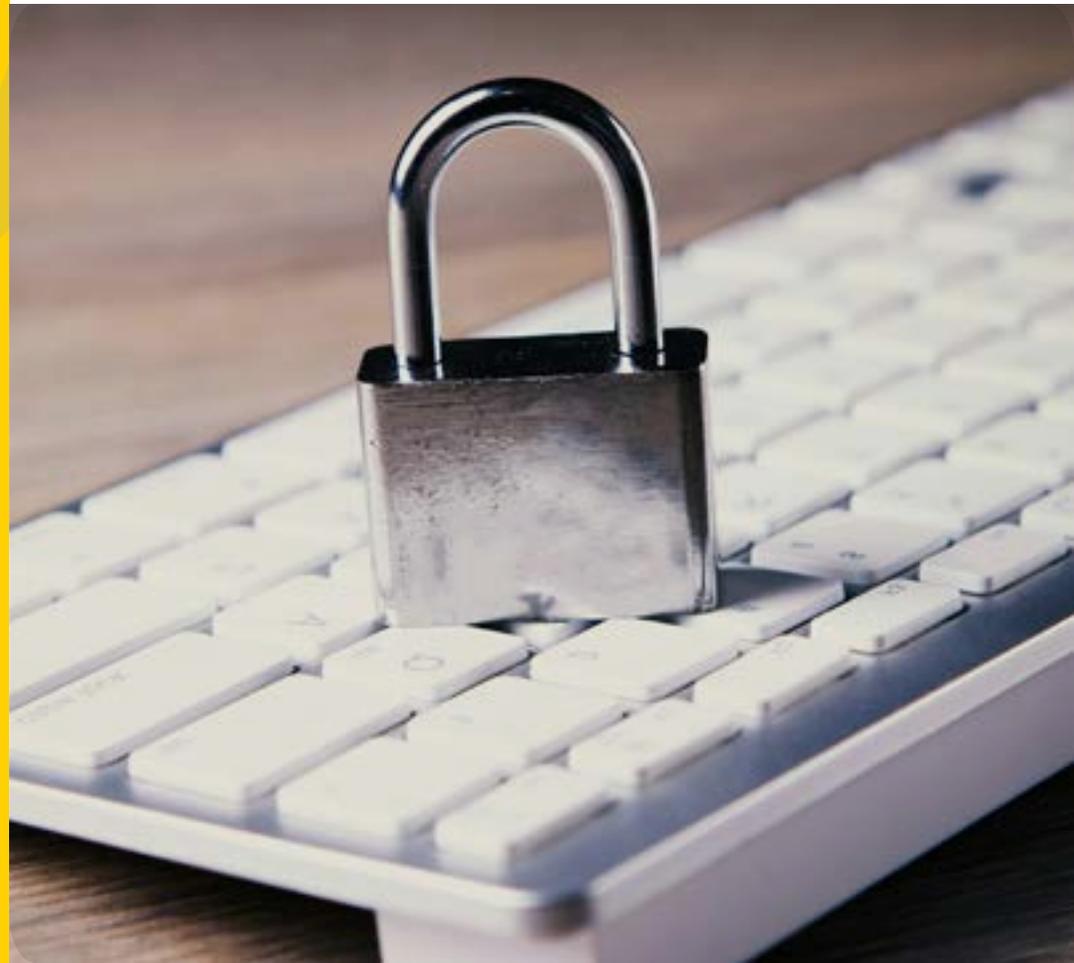


Additionally, computer vision integrates other fundamental components that together form a complete technological ecosystem: from **image segmentation** to separate objects from backgrounds; **real-time video processing** to analyze continuous frame sequences; **sensor calibration** for accurate distance and angle measurement; and **optimization algorithms** that enhance system performance and allow the system to learn from labeled data and discover patterns on its own.

From innovation to protocol: Towards smart security as a standard

Computer vision has moved beyond the emerging technology phase and has established itself as an operational tool in the field security and beyond. Despite the lingering uncertainty surrounding the ethical and legal issues underlying its use, according to the theory of “technological domestication” (Silverstone, Hirsch, and Morley, 1992), it can be considered that this **technology has been practically domesticated**, having moved beyond the integration phase to establish itself as an innovation adapted to the market reality in organization and users.

This phenomenon is clearly observed in the **rise of intelligent surveillance systems**, as what began as isolated experiments in video analysis and pattern recognition is now part of the standard in security arsenals —from automatic intruder detection to biometric access control— solidifying computer vision as a common component in security services.



The magnitude of this shift and the widespread adoption of the technology are evident in market projections, which estimate that the global artificial intelligence computer vision market will **grow from \$56.4 billion in 2025 to \$117 billion by 2030**, driven largely by the rise of smart surveillance and autonomous vehicles. Part of the technological and cultural capital of organizations, it stands as a resource that contributes to defining new operating protocols and strengthening the resilience of protection systems. In synthesis, the

domestication of computer vision signals its maturity: this technology is no longer a marginal experiment, but an established component in the operational fabric — especially in security— that manifests itself as an assumed competitive value.



04

**Integrating
computer
vision:
Technological
process and
human sensitivity**

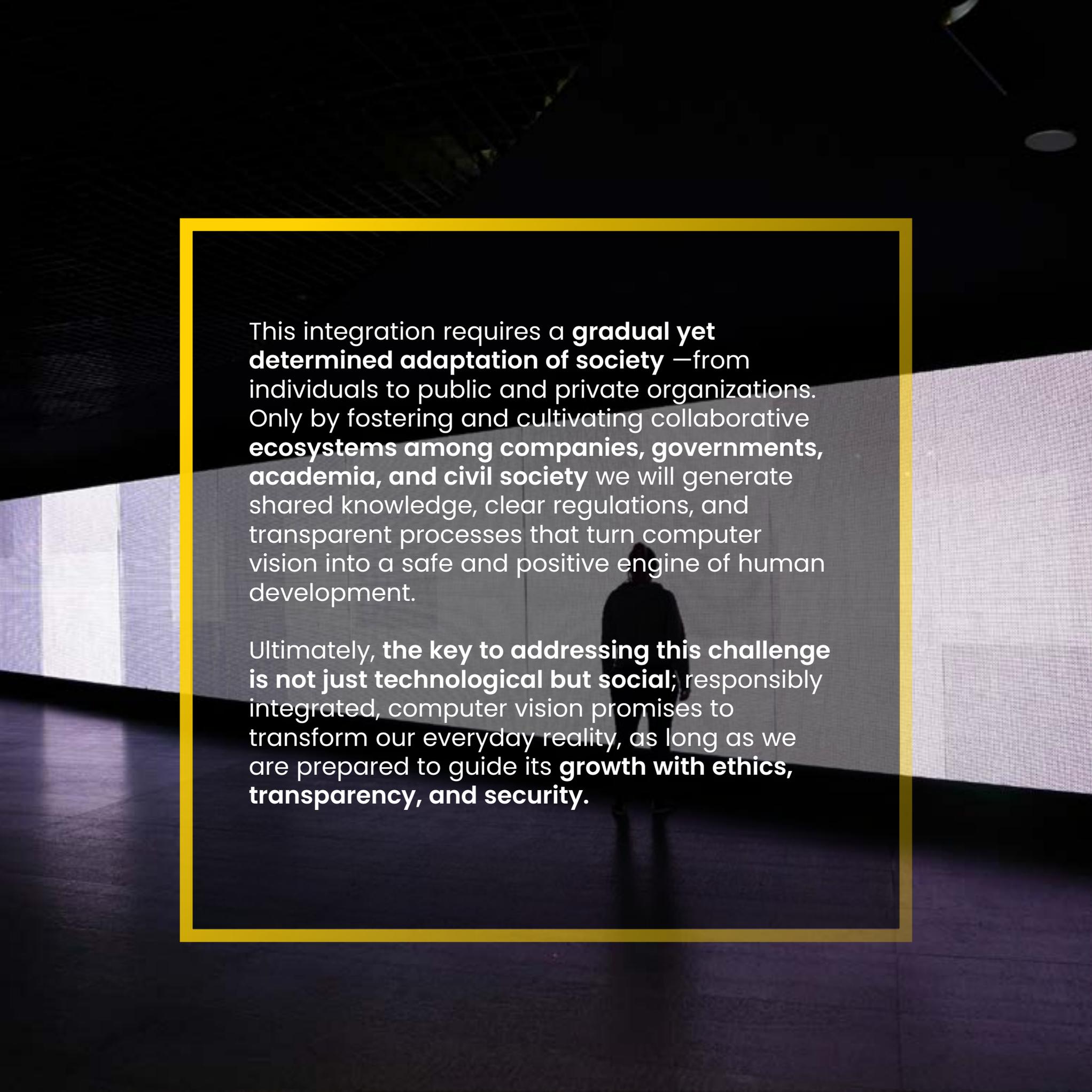


The real challenge of computer vision lies not only in its technological development, but also in **how society manages to integrate it effectively, safely, and responsibly** into the essential tools of daily life. As discussed, this technology represents a disruptive innovation in key sectors like security, healthcare, mobility, and industry, requiring **not only technological adaptation, but also cultural, ethical, and regulatory transformation.**

Effective implementation involves incorporating computer vision into **critical tools** already embedded in the **social and corporate fabric**: surveillance cameras, medical diagnostic systems, autonomous vehicles, industrial controls, and identification and authentication systems. This integration must be **progressive**, ensuring that the adoption of this technology **preserves privacy and respects fundamental rights**, as an essential balance between **progress and protection** in the hybrid security paradigm promoted by Prosegur, combining the power of artificial intelligence with human sensitivity and judgment.

At the same time, it is essential to promote **clear and accessible communication** so that it allows all social actors to understand the implications of this technological revolution, along with regulations that align with the technical progress we envision for the future of civilization. It is therefore a question of **educating, informing, and raising public awareness among citizens** to show how this technology can substantially improve their quality of life.



A person is walking away from the camera in a dark, modern interior space. The walls are illuminated with large, rectangular light projections that create a grid-like pattern. The floor is dark and reflective. The overall atmosphere is futuristic and high-tech.

This integration requires a **gradual yet determined adaptation of society** —from individuals to public and private organizations. Only by fostering and cultivating collaborative **ecosystems among companies, governments, academia, and civil society** we will generate shared knowledge, clear regulations, and transparent processes that turn computer vision into a safe and positive engine of human development.

Ultimately, **the key to addressing this challenge is not just technological but social**; responsibly integrated, computer vision promises to transform our everyday reality, as long as we are prepared to guide its **growth with ethics, transparency, and security.**



PROSEGUR
SECURITY

We guarantee the safety
of people, companies
and society as a whole.

research@prosegur.com

PROSEGUR RESEARCH

www.prosegurresearch.com