2024 | Whitepaper

OpenVINO

TensorFlow

ISO.

Viso Suite The One No Code Computer Vision Platform

Computer vision is highly disruptive across all industries. To enable computer vision in real-world applications, organizations have to manage the entire lifecycle of applications. A unified platform is needed to integrate, optimize, and accelerate AI vision adoption.

Commercial Value: Companies in various industries have started to implement a growing number of computer vision applications to save costs, improve efficiency, increase revenues, and innovate. Seeing a surge of applications, the computer vision market is expected to grow to USD 51.3 billion by 2026 at a compound annual growth rate (CAGR) of 26.3 percent. Industry analysts estimate that organizations will invest more than \$250 billion annually in digital transformation software by 2025. According to McKinsey, companies will generate more than \$20 trillion annually in added value from the use of these new technologies. This is the fastest-growing enterprise software market in history.

Computer Vision is nontrivial: The problems that have to be addressed to enable today's computer vision are highly complex. Deep learning is a major breakthrough that has dramatically transformed modern computer vision and opened the door to implementing large-scale AI vision applications. Robust cloud and edge infrastructures, a range of data services, and IoT communication are prerequisites to develop, provision, and operate computer vision applications at scale.

Most organizations will eventually need multiple and highly integrated computer vision applications. To avoid sunk costs and loss of investment, they need an efficient computer vision architecture to streamline and accelerate the development throughout the entire lifecycle. Let's start with looking into the requirements for such a platform.

"Viso Suite allows to take video streams, analyze them in a deep-learning model and draw insights in real-time at the edge, which results in real business impact."

- Arkadiusz Hruszowiec, Global Sales Director, Intel Inc

Reference Computer Vision Platform Architecture

The problems that have to be addressed to enable today's computer vision are very complex. The array of capabilities necessary for computer vision applications presents multiple development problems. Each of these represents a problem on the order of magnitude of a relatively simple software application like a CRM. This is a highly complex problem. Let's take a look at some of these requirements.

Edge AI Services: The enormous amount of data and the need for high performance, low latency, resilience, and data privacy require Edge AI capabilities. Edge AI is key to delivering computer vision at scale by moving AI workloads to the network edge. Such architecture needs optimized edge-cloud infrastructure, edge gateways, edge device management, and IoT communication.

Image Data Integration: The basis for Machine Learning at an industrial scale is the availability of a unified approach to integrating a multitude of cameras and live streams, from network cameras to webcams and special cameras.

Data Persistence: The data aggregated and processed in computer vision includes massive amounts of video and metrics data. A multiplicity of database technologies is needed, including time series, relational, big data, NoSQL, object storage and distributed file systems.

Platform Services: A myriad of sophisticated platform services are necessary for any enterprise computer vision application. Examples include access control, encryption of data in motion and at rest, load-balancing, queuing, pipeline management, multitenancy, authentication, authorization, cybersecurity, autoscaling, and data privacy.

Computer Vision: The platform needs to support a wide range of traditional computer vision and deep learning tasks, including real-time object detection, classification, segmentation, tracking, movement recognition, face recognition, pose estimation, and keypoint tracking. The highly divergent image processing requirements include continuous analytics processing, parallel and batch processing, and stream processing. **Image Processing:** Vision pipelines require capabilities to pre-process frames. Image processing includes cropping, blurring, sharpening, de-noising, feature extraction, and color and pattern recognition.

Machine Learning Services: The whole point is to enable data scientists to continuously develop and deploy Machine Learning models. Therefore, a range of tools is needed, including Jupyter Notebooks, Python, JavaScript, and Docker integrations. Increasingly important is an extensible curation of Machine Learning libraries such as TensorFlow, Caffe, Torch, or OpenVINO. An effective computer vision platform needs to support them all.

Data Visualization Tools: Any viable AI architecture needs to enable powerful data visualization tools and data connectors to integrate Grafana, Tableau, Power BI, Oracle BI, Excel, and others.

Developer Tools: Organizations require support for professional developer tools, an integrated IDE, terminal access, docker integration, and addon SDKs – or the platform solution will be rejected as unusable by the IT development and computer vision teams.

Open, Extensible, Future-Proof: The current pace of software and algorithm innovation is blinding. The techniques used today will be obsolete in 5 to 10 years. Therefore, a computer vision platform architecture must provide the capability to replace any components with their next-generation improvements. Moreover, the platform must enable the incorporation of any new open source or proprietary software innovations without adversely affecting the functionality or performance of an organization's existing applications. This is a level-zero requirement.

To meet this extensive set of requirements, viso.ai has spent years of developing and enhancing Viso Suite. The Viso Suite has been refined, tested, and proven in some of the most demanding industries and production environments – including electric utilities, manufacturing, healthcare, telecommunications, and public sector.

Challenges – Computer Vision Toolchains

Traditional Computer Vision solutions require integrating multiple tools, cameras, databases, libraries, frameworks, and platforms to build a computer vision toolchain.

Navigating the sea of tools: Today, the market is awash in open source "AI Platforms" that present themselves as comprehensive solutions sufficient to deliver Computer Vision. They all appear to do the same thing, and they all seem to provide a complete computer vision platform. While many of these products are useful, the simple fact is that none offer the scope of features necessary to develop, scale and operate enterprise computer vision applications. TensorFlow, for example, a set of math libraries published by Google, is useful in enabling certain types of ML models. For that purpose, it is an effective product. But that functionality represents perhaps one percent of an overall solution.

Stitching tools together: All these utilities are useful to solve parts of the computer vision lifecycle, but none is sufficient by itself. Each addresses only a small part of the problem required to deliver computer vision applications. Therefore, companies need to stitch different tools together in an attempt to build a computer vision toolchain. However, these utilities are written in different languages, with different computational models, with frequently incompatible data structures, and were never designed to work together.

Do it yourself: Larger organizations tend to attempt to develop a standardized approach to deliver complex computer vision applications in a "do it yourself" approach. They spend hundreds of person-years and hundreds of millions of dollars on those projects, often using open-source software with a combination of cloud products from AWS, Google, and Azure, organizing tools into the reference computer vision platform architecture depicted above. As outlined, the complexity of such a system is a multitude greater than developing a CRM or ERP system. In the past, there have been many attempts to build a CRM in-house, and as we see today, the vast majority were not successful.





"If you want to speed up the delivery cycle of computer vision, you need to simplify your ecosystem. And we've been doing that with Viso Suite along the way. It's critical for developers to have one single point of contact and one simple, no-code interface to increase the speed of delivery."

- Francesco Traversa, Data Scientist, A2A

1. Complexity

Using structured programming, the number of software API connections that one needs to establish, harden, test, and verify for a complex AI vision system is enormous. The developers of the system need to individually and collectively understand that level of complexity to get it to work. The number of engineers capable of dealing with that level of complexity is quite small.

Aside from the platform developers, the application developers and business teams also need to understand the complexity of the architecture and all the underlying data and process dependencies in order to develop any application. As business requirements evolve, the level of complexity further increases, setting many projects up for failure. This is why computer vision projects often fail to move beyond the PoC phase when requirements are adjusted to production-grade standards.

2. Robustness

Spaghetti-code applications of this nature are highly dependent upon each and every component working properly. If one developer introduces a bug into any one of the components, all applications may cease to function. Replacing tools, ML models, or functions may impact all computer vision applications.

3. Future Proof

As new libraries, new tools, and new Machine Learning models become available, those new utilities need to be available within the platform. Being limited to specific hardware, cameras, frameworks, or tools may prevent benefitting from newer, better, and more cost-efficient technologies. Consequently, every application that was built on the platform will likely need to be reprogrammed in order to function correctly. This may take months to years.

4. Data Integration

Data integration across all tools is absolutely necessary for computer vision. Using structured programming and API-driven architecture makes it very difficult to develop an integrated data architecture for multiple and large-scale computer vision applications. This is the main reason tens to hundreds of millions of dollars get spent, and several years later, no applications are deployed.

5. Toolchain Tax

Building a computer vision toolchain with different products and tools leads to having to pay for hidden costs. Such costs include the time and cost to procure the different point tools, to integrate all of these tools, and to train users on all of the tools individually. Additional hidden costs include increased security risks, with limited ability to audit and govern all activities throughout the toolchain.

6. Collaboration Gaps

Computer Vision requires cross-functional teams to collaborate effectively and keep all activities along the computer vision toolchain aligned – from choosing ML models to application design, data model, application deployment, and continuous optimization and updating. Separated tools hinder the ability to establish effective collaboration and may slow time to market, increase security risks, and result in applications that do not meet business requirements.

"In Computer Vision, it's a hassle to get the video, store it, run the deep learning model, and use the output data. Viso Suite's visual programming interface greatly simplifies and accelerates the entire process."

- Mihai Cvasnievschi, Al Engineer, Stefanini Group

Solution – Single No-code Platform Architecture

Viso Suite is a complete Computer Vision platform designed from the ground up as a single, end-to-end architecture. From data collection and AI model training to application development, deployment, and monitoring – Viso Suite's capabilities are built-in as part of the platform, so you don't have to integrate multiple tools.

ලි	IJ	Œ	2	<u> </u>	200	<u>88</u>	Ś	\bigtriangledown
Collect	Annotate	Train	Develop	Deploy	Connect	Monitor	Maintain	Secure
Data Collection	Video Annotation	Al Model Training	No Code Applications	Device Management	Edge Gateway	Event Monitoring	Continuous Delivery	Zero Trust Security
Video Capture	Dataset Manager	Neural Networks	Computer Vision Flows	Edge Fleets	Distributed Edge	Collector Agent	System Alerts	Edge Defender Scans
Data Manager	Automatic	Al Model	Visual Editor	Edge Device Virtualization	Private Edge Al Processing	Real Time Analytics	Updates (OTA)	Event Logging
Automated Collection	Labeling	Management	App Library	Profiles	Edge Tunneling	Time Series Services	Remote Health Diagnosis	Multi-Layer
Data Import	Third-Party Annotators	Version Management	Cross-Platform Hardware	Automated Deployment	(VPN)	BI Dashboard Builder	Remote Troubleshooting	RBAC
External Databases	Annotation Metrics	Cross ML Platforms	Version Management	Configuration Manager	Cloud-Edge Security	Edge Hardware Metrics	Zero Trust SSH Access	Authorization
Camera Connectors	Multi-tenant	Al Model Zoo	Containers	Edge Load Balancing	Message	Connector	Disaster	MFA
		Protection	module obre	Cloud Twin	Broker		Recovery	
				No-Code / Low	-Code			
			Clo	oud-Edge Infras	structure			

Viso Suite

Abstraction Layer: Viso Suite's model-driven architecture provides an abstraction layer that vastly simplifies and accelerates the development and deployment of computer vision applications throughout the entire lifecycle. The model-driven architecture provides a powerful solution to overcome the limitations of structured programming for highly complex problems like computer vision applications.

Integration: By greatly simplifying integration and development tasks, the application developer does not have to be concerned with all the data types, data interconnections, and processes that act on the data. Instead, the user simply needs to link modules, and all the underlying data, APIs, associations, connections, and processes associated with or used to manipulate those data are abstracted in the model itself. This dramatically

reduces the number of elements, processes, and connections of which the user needs to be aware, making computer vision pipelines tractable, scalable, and agile.

No-Code and Low-Code: The Viso Suite provides unique no-code capabilities throughout the computer vision lifecycle to reduce complexity and increase productivity. This frees developers from having to worry about writing code for data mapping, hardware management, edge-cloud infrastructure, or the mechanics of the myriad of computational processes like buffering, streaming, load-balancing, queuing, pipeline management, encryption, and so on. Built-in low-code capabilities allow adding custom code, docker containers, modules, datasets, or Al models to extend every aspect of the platform.



Model-driven Development: The optimal design for a computer vision application uses abstract models as placeholders to which a developer can link an appropriate function. Using Viso Suite, developers create computer vision pipelines by wiring together cameras, processors, deep learning models, and image processing functions. For example, an application might include a module called Object Detection, that in turn serves as a placeholder that might incorporate any Deep Learning Framework like TensorFlow or OpenVINO and any model such as YOLO, SSD, or ResNet. The AI workloads might be computed with processors like CPU, GPU, VPU, or TPU. A video input module might contain several cameras of different types, external videos, or virtual cameras. And so on.

Future Proofed: Another important capability of the Viso Suite enabled by its model-driven architecture is that the applications developed on the platform are entirely futureproofed. Suppose, for example, that an organization developed all its applications using YOLOv3 as the object detection model and then later decided to switch to an alternate algorithm – such as YOLOv7 that runs much faster on several times cheaper hardware. All applications deployed previously using YOLOv3 will continue to run without modification after that replacement. Similarly, computing hardware, cameras, and image annotation or model training environments can be replaced without disrupting current setups. This enables organizations to immediately and easily take advantage of new and improved product offerings as they become available. **Platform independence:** The advances in computer vision over the past years have been dramatic and continue to accelerate. While corporate leaders eagerly embrace AI adoption, they are also very concerned about vendor lockin. They want to be able to negotiate continually, deploy different applications to different environments from different vendors, and they want to be free to move applications from one platform to another. Therefore, the ability to deploy anywhere, using any camera sensor and ML framework, is an additional requirement of modern computer vision platforms that is fully supported by the Viso Suite. Applications developed with the Viso Suite can run on various platforms of any vendor, on the edge, virtual edge, virtual servers in the cloud, MEC, and behind the firewall in corporate networks.

Scalability: Viso Suite enables organizations to build and maintain a portfolio of computer vision applications. Automated infrastructure provides horizontal scaling to a fleet of edge devices capable of processing thousands of camera streams. The model-driven architecture of the Viso Suite enables unique capabilities to remotely configure, diagnose, and troubleshoot endpoints effectively, helping operations teams to significantly facilitate the usually difficult and expensive maintenance of large-scale vision systems.

"We've built computer vision applications using Viso Suite. It allows us to deliver business value faster and easier."

- Carlos Arredondo, Al Lab Lead, PwC

1. Reduce complexity and lifecycle costs

By dramatically reducing the number of entities, objects, and processes the developer needs to understand and by freeing the developer from constantly keeping up with the technicalities, the no-code architecture of Viso Suite decreases the cost and complexity of designing, training, developing, testing, provisioning, maintaining, and operating computer vision by as much as 100 times or more.

2. Accelerate time to results

Today, developing real-world computer vision solutions with heterogeneous toolchains is inflexible and slow, with rigid technical requirements and little room for changes. The Viso Suite leverages a model-driven architecture with powerful no-code and low-code interfaces throughout the computer vision lifecycle. As a result, custom applications can be developed and delivered dramatically faster, in days instead of months – for every recurring development cycle.

3. Reduce risk with a unified security model

Having to manage authentication and authorization across many tools is painful. Viso Suite provides users with a single login and one place to set permissions and password policies, so everyone has the correct access. Viso Suite provides industry best practices for highsecurity encryption, secure communication, and zero-trust security.

4. Reduce administrative overhead costs

With a single solution to implement for covering the entire computer vision toolchain and delivering all computer vision applications, there is a significant reduction of administrative overhead. Smaller teams are needed to administer a single solution versus a complex toolchain. Also, engineers only need to be trained on using one tool instead of dozens. As a result, more engineering resources can be allocated toward developing computer vision capabilities.

5. Reduce application downtime

Diagnosing and troubleshooting issues in computer vision deployments is very challenging. With traditional toolchains, teams likely don't all have access to the same applications and need to move tickets back and forth trying to troubleshoot issues. Using Viso Suite, developers can quickly spot issues with automated health scans, directly troubleshoot, track changes, and roll back to previous versions. With everyone on the same page, and unique remote capabilities, troubleshooting is much simpler, and things get up and running faster.

6. Unify all computer vision applications

In the early stages of computer vision adoption, companies often implement point solutions. Eventually, most organizations need multiple computer vision applications, require customization, integration into IT systems, and a level of flexibility and control that point solutions cannot provide. As an integrated enterprise platform, Viso Suite provides powerful capabilities to build and scale any computer vision application while leveraging strong synergies and learning effects.

7. Implement privacy and full control

Private computer vision is imperative for any kind of business application. Sending video for inference to the cloud or sharing data and AI models between tools pose enormous risks. Viso Suite was built to power private Edge AI processing, with on-device inferencing that does not require storing or sending video data. By integrating the entire toolchain, it becomes possible to fully control and protect all data, data processing, AI models, and applications.

"Viso Suite provided a way to build and rollout a flexible AI vision system to production, achieving significant cost savings."

- Thomas J. Ernst, Fleet Manager, Swiss Post

Capabilities of the One Computer Vision Platform

1. Collect

Collect data for computer vision annotation with Viso Suite. Use powerful, automated collection capabilities to gather high-quality training data using any camera enrolled in Viso Suite. Use continuous data collection to improve the performance of AI models.

2. Annotate

Annotate image and video data with Viso Suite. Bring teams together to create high-quality datasets using a complete palette of labeling tools. Avoid exposing your data by transferring it to third-party tools and services. Your data never leaves your workspace, keep everything safe and stay in full control.

3. Train

Manage AI models of all frameworks in one place. Make them available to be used in applications. Import AI models by choosing from pre-trained deep learning models, or import your own. Train custom AI models in modular environments. Manage versions easily, constantly update, improve and re-train your AI models.

4. Develop

Develop computer vision pipelines with the intuitive no-code editor. Create custom applications with modular building-blocks. Wire together cameras, AI models, image processing, and computing hardware. You will never hit a wall. Extend everything with custom code, containers, or modules. Continuously evolve and update applications with built-in version control.

5. Deploy

Deploy computer vision applications at scale. Viso Suite provides a complete device management. Enroll and manage a fleet of devices safely, without coding or manual installation. Viso provides a fully automated deployment manager. Use profiles to efficiently roll out new releases and manage different deployment environments. Now you can ensure your applications scale.

viso.ai







6. Connect

Connect deployed applications and devices. Process all video streams ondevice with private Edge AI processing, and avoid sending video data to the Cloud. The IoT Edge Gateway manages all Edge-Cloud connectivity with secure edge tunneling (VPN) and multi-layered security and encryption. The automated collector agent provides scalable IoT communication, offline capabilities, local metrics aggregation, and pre-processing.

7. Monitor

Monitor events and metrics in the Cloud. Viso Suite provides real-time analytics with time-series databases. Monitor detailed hardware metrics at the device level. Build custom BI dashboards and reports to visualize application metrics with ready-made charts. Explore and drill-down historical data, and create custom queries. Integrate your BI systems with the data connector.

8. Maintain

Maintain all applications and ensure continuous delivery. The automated infrastructure and no-code capabilities of Viso Suite enable the implementation of a responsive and fast computer vision development, update and upkeep strategy. Detect issues early with powerful diagnosis tools, system alerts, and automated health checks. When issues are identified, Viso Suite provides integrated tools for remote troubleshooting, fast roll-backs, and disaster recovery.

9. Secure

Secure and protect all data and applications. Viso Suite deeply integrates enterprise-grade security capabilities for the entire computer vision lifecycle. Viso Zero Trust Security provides multiple layers of protection. Manage teams with a powerful user management, role-based access control (RBAC), and custom user roles. Implement user password policies, session management, and multi-factor authentication (MFA). Log and trace all activities and events.

Deliver secure applications with dependency scanning and container vulnerability scans. Protect data with high-security algorithms, AES-256 for data encryption at rest, SHA-256 and RSA-2048 for data in transit. Cloud workspaces run in secure virtual private clouds (VPC), and cloud-edge communication uses secure virtual private networks (VPN). Edge devices are fully locked down, with cloud authentication and device twin. The Edge Defender provides container whitelisting and auto wiping. Secure remote edge access with SSH Zero Trust Security.











Conclusion – A Powerful, Unique Computer Vision Platform

The Viso Suite provides a uniquely powerful platform enabling organizations across industries to develop and operate enterprise computer vision and deep learning applications at scale, with a fraction of the effort and resources required by other approaches. The model-driven approach to developing computer vision applications using the Viso Suite has been tested and proven in dozens of highly demanding and customized real-world deployments at some of the world's leading organizations. The model-driven and no-code development increases productivity by maximizing compatibility between systems, simplifying the process of design, promoting effective communication between developers and teams, and establishing standardization across a portfolio of different computer vision applications.

Applications built with the Viso Suite are flexible, easily upgraded, and can be ported across different hardware platforms with little or no modification, providing a solution that future-proofs customers' investment in computer vision application development.

About viso.ai

Swiss 🛟 Made

viso.ai[®] is a leading computer vision infrastructure provider. We are backed by Accel, one of the world's largest VC investors that has invested in Facebook, Slack, Atlassian, Spotify, Venmo, Celonis, and Dropbox.

viso.ai delivers the Viso Suite for developing, deploying, and operating large-scale computer vision applications. The core of the viso.ai offering is a revolutionary, no-code driven AI architecture that dramatically enhances and accelerates the delivery of computer vision. Operating out of Switzerland, viso.ai serves Fortune 100, Global Fortune 2000, and Governmental Organizations worldwide, and is a technology partner of Intel, Nvidia, HP Enterprise, Amazon Web Services, DXC Technology and Samsung.