

# Develop Computer Vision AI Quickly and Easily with Intel® Geti™ Platform

## Intel's computer vision AI platform for data labeling, training, optimization, and retraining

### Key Takeaways

- Intel Geti platform enables the rapid development of custom models by reducing labor-intensive tasks and harnessing greater collaboration between teams.
- Intel Geti platform works with most AI frameworks and exports the trained models in native TensorFlow and PyTorch formats, plus OpenVINO™ toolkit intermediate representation (IR) files that are optimized for Intel® hardware.
- REST APIs and the software development kit (SDK) enable users to integrate the Intel Geti platform into their value chain to push data directly into the platform and pull trained models directly into their deployment pipelines.

### The Challenge: Putting AI to Work Across Industries

AI is making major strides in research labs where computing power, time, and technical expertise are a given. In the lab, data is clean, experiments are controlled, and deep learning models can go through many iterations to reach accurate results. Unfortunately, enterprise teams have limited resources and time for experimentation. If AI is going to solve practical problems, implementing it has to be quick, reliable, and far easier to do.

### A New Solution: Intel® Geti™ Platform

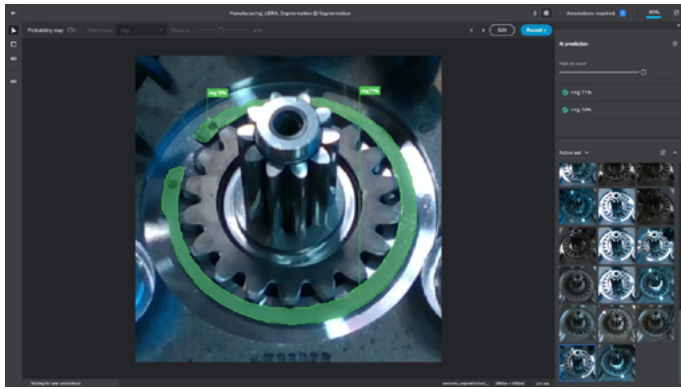
Intel is working across several fronts to put AI to work in industrial and commercial enterprise applications. One of the key bottlenecks is the technical difficulty of training AI to do a new task. Traditional training can take weeks, even months, of data science resources and immense computing power. These challenges make AI model development resource and time prohibitive for many applications.

To break these bottlenecks, Intel has created the new Intel® Geti™ computer vision platform—a collaborative, intuitive platform where data scientists, developers, and domain experts can work together to train deep learning models for specific computer vision applications. The Intel Geti platform speeds up model development by simplifying labor-intensive tasks and harnessing greater collaboration between teams on one single platform for data labeling, model training, optimization, and retraining. Most importantly, the solution unlocks faster time-to-value for digitalization initiatives with AI.

### Intel Geti Platform's Active Learning Workflow for Training

The Intel Geti platform uses a type of machine learning called active learning to train convolutional neural network (CNN) models. With active learning, human experts stay in the loop and teach the machine as it trains. Active learning picks the most useful data for users to tag and reduces the number of annotations needed as well as the sample bias. This helps neural network models learn fast from the most informative data samples, thereby making data labeling and model training fast and less labor intensive.

1. Upload a data set—20 to 30 images to start.
2. After the first round of training, the Intel Geti platform reviews the data and chooses the best candidates for training.
3. Continue to select what you want the platform to find and label it.
4. The Intel Geti platform begins retraining the model and making predictions.
5. Accept or correct predictions and label more data.
6. The model's accuracy improves rapidly.
7. Export runtime models with a click.
8. To retrain, upload new images, add them to the data set, and export an updated model.

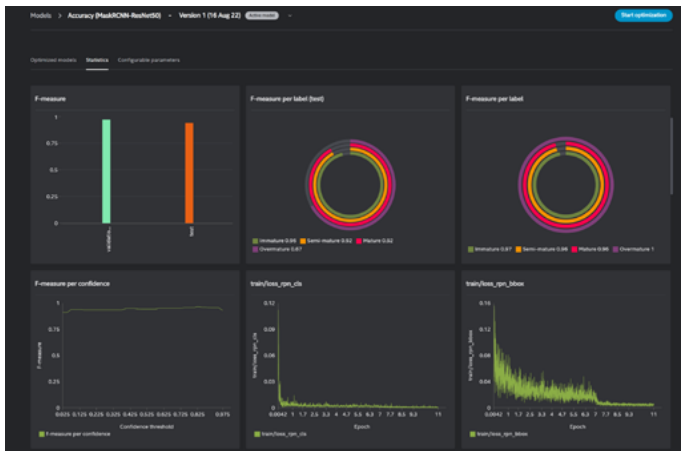


**Figure 1.** The Intel Geti platform uses active learning to simplify labeling and speed up AI training<sup>1</sup>

**Automate Tasks with an Intuitive, Time-Saving Platform**

The Intel Geti platform includes the kind of time-saving capabilities you would expect in a professional image-editing application, applied to common data-science and image-processing tasks. It has features for knocking out backgrounds, autodetecting similar objects, and segmenting shapes. The Intel Geti platform makes batch processing, image analysis, and labeling jobs drag-and-drop simple.

- Drawing tools for polygons, rectangular, rotated, and circular bounding boxes
- CV annotation assistants for grab cuts and watershed transformations
- Object detection assistants that identify similar shaped objects automatically



**Figure 2.** The Intel Geti platform presents telemetry on models in easy-to-read data

**Gather Training Feedback from Multiple Experts—Even Remote Staff**

Want to get training input from radiologists throughout a healthcare system? Need input from all the quality control engineers at an automotive manufacturer? With the Intel Geti platform, you can collaborate and gather training input from hundreds of experts and train models remotely—either

in the cloud or on-premises data center. This distributed approach to training gathers vast amounts of knowledge fast and can help reduce sampling bias by incorporating multiple experts’ viewpoints into the model.

**Fine-tune Parameters with the Same GUI**

The Intel Geti platform extends user-friendly features to advanced data science training parameters. Teams stay in complete control of fine details without dropping into Python. Plus, they can see the result of fine-tuning immediately in the model’s performance.

- Set learning parameters, including rate, batch size, and number of epochs
- Define active learning parameters like number of annotations required for initial training, making predictions, and retraining
- Adjust parameters for rotation, scaling, and other augmentations
- Define pre- and postprocessing for image sizes, soft thresholds, blur strength

**Export Models and Inference Packages for Practically Any Hardware**

The Intel Geti platform exports production-ready deep learning models in several formats, including Google’s open source TensorFlow framework, PyTorch, and open source OpenVINO™ toolkit files (intermediate representation [IR] files with an .xml and .bin file) for Intel® CPUs, GPUs, and VPUs. The models you train in the Intel Geti platform will run on virtually any laptop, desktop, embedded PC, or server—including the machines you already have.

**Retrain with Production Data Anytime**

Automating labeling and training is a major time-saver, but it doesn’t change the experimental nature of getting deep learning models to work effectively in challenging enterprise applications. They may respond inaccurately to new scenarios or drift over time.

With the Intel Geti platform, teams can use the same incremental, active learning method to retrain the model. Just upload new production images and add them to the existing model. Then label and train until the model gains the accuracy required. When it is done, export the model for further testing or production.

**Optimize Runtime Performance with the Intel® Distribution of OpenVINO™ Toolkit**

The OpenVINO toolkit has several post-training optimization tools such as NNCF (Neural Network Compression Framework) for compressing and quantizing the model during and after training or POT for post-training quantization. By shifting precision from FP32 to FP16 or quantizing to INT8, you can often reduce the model’s file size and compute needs and increase throughput with minimal losses in accuracy. Running at INT8 also increases performance on Intel CPUs with Intel® Deep Learning Boost, a set of Intel® AVX-512 instructions that accelerate inference calculations.

## Get a Head Start on Application Development with Intel® Tools for AI Inference

Getting to a trained model is just one step on the way to a production application. Use the [OpenVINO toolkit](#) to build analytics pipelines, optimize performance, and package models with inference engines that will run on any mix of Intel hardware. You can run inference packages with stand-alone applications or deploy them as containers using open source or commercial platforms.

- [OpenVINO Model Server](#), an open-source high-performance system for serving machine learning models, with the same architecture as that of TensorFlow Serving API
- [Kubeflow](#), an open-source machine learning toolkit for Kubernetes
- [OpenVINO toolkit operator](#) for [Red Hat OpenShift Data Science](#) and its open source counterpart [Open Data Hub](#), which are AI/ML platforms for hybrid, multi-cloud workflows and deployments

Intel also provides building blocks for deep learning application development, including middleware for time-series and video analytics, plus reference implementations for industrial, medical, retail, and smart city use cases. You can find them all at the [Intel® Edge Software Hub](#).

## Intel Geti Platform Use Cases

The Intel Geti platform retrains convolutional neural network models for key computer vision tasks, including:

- Axis-aligned as well as rotated object detection
- Semantic as well as instance segmentation, including counting
- Single-label, multi-label, and hierarchical classification
- Anomaly classification, detection, and segmentation

It also supports task chaining, so you can build multistep, smart applications.

### Manufacturing

Develop AI for autonomous assembly, defect detection, industrial controls, and worker safety systems.

### Smart Agriculture

Develop models for autonomous machines to analyze crop health; identify pests and weeds, then spot-fertilize and treat; and harvest crops.

### Smart Cities

Develop AI for traffic-management systems to route traffic automatically, build solutions that recognize and respond to emergencies, and use camera data to improve safety in near-real time.

### Retail

Develop AI for autonomous inventory management systems; smarter safety and loss prevention; and accurate, touchless checkout.

### Video Safety

Develop task-specific models for video analytics, social distancing, PPE, and safety gear detection.

### Healthcare

Develop models to detect anomalies in medical images, assist in diagnostics and procedures, analyze lab results and count cultures, and speed medical research.

All these models can be deployed with OpenVINO toolkit or can be integrated into your pipeline with REST APIs and a software development kit (SDK).

## Intel Geti Platform Success Stories

These early adopters are using the Intel Geti platform to train deep learning models for complex visual tasks, invent new uses for smart computer vision, and reduce bias by expanding the AI training process to more experts.

### Disease Recognition: Royal Brompton and Harefield Hospitals, London



Medical specialists are training models to review electron microscope images and transform their research.

### Agricultural AI: Small Robot Company



Small Robot Company is using the Intel Geti platform to train its robots to identify and monitor each plant in a farmer's fields for crop health, weed identification, and pest control.

### Biodiversity Research: Naturalis Biodiversity Center



Researchers are creating deep learning models for a network of cameras that detect insects, enabling them to take big steps forward in biodiversity monitoring.

### Smart Cities: Sensing Feeling



Sensing Feeling is replacing its use of off-the-shelf AI models with custom models built using the Intel Geti platform. These custom models enable the classification of pedestrians, cyclists, e-scooters, and users of other forms of micro-mobility in urban environments with the aim of improving infrastructure.

### Manufacturing Quality Control: Bravent



IT consultancy Bravent is developing an efficient, flexible AI-based computer vision solution to reduce human error in complex machinery assemblies, which can be scaled and adapted across a range of industries.

## Intel Geti Platform Technical Details

### A Vital Link in the Complete Intel® AI Portfolio

Intel® AI tools for data science and AI developers range from the Intel® oneAPI Analytics Toolkit and optimized distributions of Apache Spark and TensorFlow to pretrained models and reference applications for specific edge use cases. The Intel Geti platform fills a missing link in the AI technical value chain—fast, effective training for deep learning models.

### Supported Tasks

- Object detection, classification, segmentation
- Anomaly classification
- Task chaining for building models with multiple analytical steps

### Supported Deep Learning Models

There are a range of deep learning model architectures supported in the Intel Geti platform today, and support for additional architectures will be coming in future releases. The table below summarizes those supported models and also provides references to academic literature for readers interested in developing a deeper understanding.

Computer vision task	Task types	Model architectures supported
Image classification	Single label, multi-label, hierarchical	LinearHead x (MobileNet-V3, EfficientNet-B0)
Object detection	—	ATSS + MobileNet-V2, SSD + MobileNet-V2, YOLOX + CSPDarkNet
Instance segmentation	Counting, rotated object detection	MaskRCNN x (ResNet 50, EfficientNet-B2)
Semantic segmentation	—	Lite-HRNet
Anomaly-based tasks	Classification, detection, segmentation	STFPM, PADIM

### Supported AI Frameworks

- TensorFlow
- PyTorch

### Supported Deep Learning Formats

- Native TensorFlow
- PyTorch
- OpenVINO toolkit IR files (.xml and .bin) for Intel hardware

### Integrations

An SDK enables you to utilize REST APIs for exporting datasets, annotations, and models directly into downstream processes. This API SDK provides functionality for (1) computer vision task creation from datasets on disk, (2) project downloading (images, videos, configuration, annotations, predictions, and models), and (3) deploying a project for local inference with OpenVINO toolkit.

### Installation Options

#### Local Server

Run on a dedicated, local machine for higher speeds and more security.

#### Cloud VM

Install on a virtual machine in the cloud, without needing to manage hardware.

### Learn More About the Intel Geti Platform

If you'd like to schedule a demo, or find out more about the Intel Geti platform, please contact us.

- [geti.intel.com](https://geti.intel.com)



### Notices & Disclaimers

<sup>1</sup> Testing completed by Bravent in August 2022. Intel Geti platform pre-production version.

Performance varies by use, configuration and other factors. Learn more on the [Performance Index site](#).

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Intel does not control or audit third-party data.

You should consult other sources to evaluate accuracy.

Intel is committed to respecting human rights and avoiding complicity in human rights abuses. See Intel's Global Human Rights Principles. Intel® products and software are intended only to be used in applications that do not cause or contribute to a violation of an internationally recognized human right.