Video analytics demonstrates a fast and effective solution to incidents of missing children

Executive summary

Deep learning-based computer vision is increasingly important, both across a multitude of vertical markets and within the academic research community. In use cases ranging from personalizing retail experiences to increasing safety in smart cities, facial recognition is a critical facet in the prevalence of AI solutions impacting our everyday lives. Like any deep learning task, facial recognition presents an incredible challenge for developers and integrators, requiring the creation and deployment of complex algorithms in order to detect, extract, and match facial features. A facial detection and recognition proof of concept from Intel demonstrates—through a missing children use case—that these algorithms can be optimized and accelerated with the help of the Intel® Vision Accelerator Design with Intel® Arria® 10 FPGA, along with an end-to-end portfolio of advanced Intel® hardware and software technologies and development tools. The resulting increases in performance and accuracy deliver greater efficiency and effectiveness for AI solutions integrating facial recognition.

Challenges

Exponentially more video data is now available for analysis and there is a growing demand that this data be available in near-real time. AI capabilities, such as face and object recognition and classification via deep neural networks, can turn streaming video into a repository of intelligence that can help improve safety and enable new experiences. The ability to rapidly identify individuals is key. Demand for capabilities derived from machine learning techniques is driving the need for increased intelligent learning infrastructure that starts with visual intelligence at the point of data capture without compromising privacy. Enabling deep learning at the edge becomes very important in order to process video and photo data locally, and minimize video and photo transmissions over the network.

In order to take advantage of the opportunity, however, solution providers and developers need extensive expertise and ample time to develop and combine algorithms, as well as an infrastructure that can effectively and efficiently process potentially massive amounts of video data at the camera, a network video recorder (NVR), an edge appliance, and an on-premise or cloud server.

Intel® Vision Accelerator Design with Intel® Arria® 10 FPGA

The new Intel Vision Accelerator Design with Intel Arria 10 FPGA provides a blueprint for plug-in PCIe® accelerator cards. The design enables developers, solution providers, and end users across a breadth of industries to take advantage of vision-based inference for advanced, highly accurate, and fast visual intelligence at the edge (e.g., edge appliance or on-premise server).
Solution

The facial detection and recognition proof of concept utilizes the Intel Vision Accelerator Design with Intel Arria 10 FPGA, a blueprint for PCIe* add-in cards, along with Intel® development tools. It allows measurement of similarity between a current facial image and references in a recognition database. The solution spans the facial recognition sequence, including the processes of facial detection, feature extraction, and facial indexing and matching. A wide range of existing and emerging use cases can be supported, including safeguarding workers in industrial settings, or even finding children who are lost or missing in a city environment. With facial detection and recognition capabilities, powered by Intel® processors and Intel Vision Accelerator Designs with Intel Arria 10 FPGAs, turnaround time for identifying a missing child can be less than 20 milliseconds, and online tracking of the child can begin as soon as the child has been identified.¹

Facial detection detects the location of ears, nose, and mouth on a person’s face. An algorithm, running on top of the OpenVINO™ toolkit, adjusts for orientation and alignment so that the face is oriented properly for identification.

Feature extraction detects the details of shapes and features on each face captured by the facial detection algorithm. The feature extraction algorithm running on top of the OpenVINO toolkit, and accelerated by the Intel® FPGA, contains 32 kernels of feature detectors to extract detail features. Each face is converted to a feature vector with 1,000 point elements.

Feature matching compares detail face features to the database store. When a new feature vector is present, the high-speed database search finds the entry with the most similar feature vector and extracts the metadata associated to that entry.

Facial detection starts with video analytics at the camera, gateway, or local “edge” server and is completed by the back-end server, which matches extracted features against the database.

Key capabilities
- Facial detection
- Landmark detection
- Facial alignment
- Feature extraction
- Facial tracking
- Facial matching

The proof of concept (PoC) runs algorithms at the camera, gateway, or local “edge” server to accelerate facial detection and recognition.
Sample configuration in brief

Let’s take a closer look at the proof of concept conducted by Intel, showing how the end-to-end facial detection and recognition reference solution can be configured to an application that aims to help solve the problem of missing children.

It starts by training a facial detection and feature extraction algorithm using public and proprietary data sets. Pretrained deep learning model data (using the PVANet* topology) was specified with facial detection. Using Intel’s proprietary topologies, the algorithm taps the coordinates of the face location in the original image, performs the facial alignment process for each face, and aligns the position and orientation of the landmarks to the center of the region of interest in the image.

The facial database was created based on a preregistered face set. Note that the missing child’s or adult’s ID and data do not need to be in a preexisting database; relatives can supply the essential data in the form of video or photos. To better secure and protect privacy of missing children and their families, the database can be populated in a back-end server, with only the video analytic algorithms deployed in the edge server and responding only to the facial features that are very close to the preregistered face features in the database. This way, no video or photos are sent over the network.

How it works

Video analytics developers have the option to place the facial detection and facial recognition algorithms on smart cameras or NVRs with Intel® Arria® 10 FPGA acceleration, depending on system-level design requirements.

Because the feature extraction model is more complex than the detection module, it requires more computer resources. The Intel Arria 10 FPGA adds power-efficient acceleration for high-performance throughput. Since feature extraction detects the location and position of faces and facial features in each frame of video, with Intel Arria 10 FPGA we can achieve 100x data compression at the edge server by sending only the face feature vectors to the back end instead of sending video over the network. The extracted feature vector can be sent to the server without consuming too much network bandwidth.

Capacity is scalable: a single server with two Intel® Optane™ 3D XPoint™ NVMe* disks can handle more than 150 million database entries. In this POC with 20 million database entries, by utilizing OpenVINO, the video analytics algorithms can run at any point in the hardware configuration—at the edge appliance, on-premise server, or AWS* or Azure* cloud server.

Once facial feature vectors are obtained from the deep learning network of an NVR equipped with FPGA acceleration, the vectors can be fed for large-scale facial matching at the back-end Intel® Xeon® Scalable processor-based servers.

To better utilize the Intel Xeon Scalable processor in the back-end server, the database can be divided into many smaller sub-database of 20 million faces, and each physical core can handle the training and searching task designated to each CPU core, so that the actual identification can happen very quickly. In the database, data is presorted according to the allotted points, with the algorithms adjusting for differences in facial orientation.

Back-end server data output indicates face matches based on the feature vector(s). Depending on the single- or multithread mode, search time can be reduced from 56 to 15 milliseconds per inquiry using one server with two-socket Intel® Xeon® Platinum processor CPUs.

The end-to-end solution is ideal for the development of public safety usage models.
Key technologies

The solution takes advantage of the high performance, programmability, and flexibility of the Intel Vision Accelerator Design with Intel Arria 10 FPGA, as well as the OpenVINO toolkit, Intel® Optane™ SSD, Intel 3D XPoint, Intel Atom® processor, Intel® Core™ processor, and Intel Xeon Platinum processor (though the solution works with any Intel Xeon Scalable processor). The set of blueprints, integrated hardware and software technologies, and software development tools provide a reliable, integrated architecture enabling efficient, effective deep neural network training and inference for AI.

Feature extraction

Feature vectors

Facial matching

Scalable processor). The set of blueprints, integrated hardware and software technologies, and software development tools provide a reliable, integrated architecture enabling efficient, effective deep neural network training and inference for AI.

Back-end servers based on Intel Xeon Scalable processors and Intel® Optane™ memory can easily support hundreds of face matches per second. Rather than needing terabytes of costly memory, Intel 3D XPoint uses Intel® Memory Drive Technology to merge memory and volume in a 3D disc to increase database capacity and allow for fast search. In our example, extended memory mode is used, which allows an Intel Optane SSD to participate in a shared memory pool with DRAM at the operating system level and above, enabling bigger or more affordable memory.

OpenVINO toolkit enables deep learning on hardware accelerators and streamlined heterogeneous execution seamlessly across Intel's silicon architectures. It includes the Intel® Deep Learning Deployment Toolkit with a model optimizer and inference engine, along with optimized computer vision libraries and functions for OpenCV® and OpenVX®. With the OpenVINO toolkit, algorithms can be moved from edge to cloud servers and back again. This helps reduce the number of servers and modules needed. Requirements or algorithms can change—without impacting the developer or customer.
**Intel Vision Accelerator Design with Intel Arria 10 FPGAs**

offers exceptional performance, flexibility, and scalability for computer vision solutions, whether at the camera, NVR, on-premise server, or in the cloud. The design adapts advanced display, video, and image processing for deep learning workloads. Intel Arria 10 FPGAs achieve high-performance images-per-second at reduced power, and provide dynamic flexibility, consistent power consumption, future-proofing for custom or new workloads, and low latency.

An Intel Arria 10 FPGA is used to execute and accelerate the feature extraction algorithm. The facial alignment algorithm locates facial landmarks, applies affinity transformation, and crops the input image. A facial landmark detection algorithm finds the eyes, eyebrow, nose, and mouth. The sample feature extraction application is integrated as part of the Intel Deep Learning Deployment Toolkit.

While it may sound complex, the robust and ever-expanding array of Intel® Vision Products is designed for seamless integration and optimized for performance and efficiency, providing a flexible, effective foundation for AI innovation.

**Conclusion**

With Intel, solution providers, developers, and vertical markets can simplify AI implementations and speed time to insight. Finding missing children is a powerful example of the ways in which video data can deepen our understanding and help solve complex challenges.

**Learn More**


Find out more about Intel innovation for AI at [intel.com/ai](http://intel.com/ai).

Download the free [OpenVINO toolkit](http://intel.com/ai).