Solution Brief

Internet of Things
Touchless Technologies

Intel-powered Outdu Video & Audio Analytics for Contactless Operations

Powered by Intel® processors, Outdu's video and audio analytics solution transform self-help kiosks and touch-screen displays into contactless devices.

Table of Contents

Executive Summary 1
Challenges 1
Solution 1
Technologies Implemented 2
Benefits of Outdu's Contactless Solution 4
Use Case Examples 4
Conclusion 5

Executive Summary

Many enterprises today deploy millions of touch-screen displays and self-help kiosks to provide service to their customers. They continue to deploy more such self-service terminals to reduce manual intervention. However, the new normal has forced operating procedures across the world to minimize contact with surfaces and disinfect them at regular intervals. Touch screens are hugely affected by this new situation. Suddenly, contactless kiosks are a very attractive proposition. But what is to be done with the existing kiosks deployed in the field?

Outdu's real-time video and audio analytics solution, built on Intel® platforms, provides an elegant solution to this question. A camera deployed on kiosks feeds user inputs for video analytics in terms of gestures that are used to interact with a virtual keyboard to enter data. In some cases, the camera reads information from QR codes displayed on a user's mobile phone. Audio analytics is used to analyze spoken keywords that customers use while navigating around the contactless screen to complete data entry tasks.

Such contactless inputs from the user are converted to actual keyboard inputs or touch-screen inputs with the help of a simple lookup table. Thus, a camera and a noise-canceling microphone are effectively used to convert an existing kiosk into a contactless one, with minimal effort, and low investment and with no change in functionality.

Challenges

Touch-screen application developers expect mouse, touch, and keyboard-based event-listeners to interact with users. With the touch aspect gone, it becomes imperative that new ways to interact must be provided – both for existing systems and new ones.

Multiple form factors make the design challenges even greater; a single device of whatever form, with the relevant software should, in effect, smoothly transform into one capable of performing contactless operations.

Any change to the existing infrastructure must be easy to implement and operate. It must not sacrifice existing functionality and there must be no need to develop new interfaces and applications, which could be time consuming and expensive.

Solution

Outdu is developing video and audio analytics solutions using neural networks with Convolution Neural Network (CNN), Long Short-Term Memory (LSTM) and other architectures. These analytics are applied in diverse use cases such as customer walk-in and audience metrics, vehicle information, and driver behavior.

These analytics are already being used to gather inputs in a contactless manner at self-service touch-screen kiosks.

Outdu's solution combines a USB camera and a noise-canceling microphone with the existing hardware (Intel's CPU with an integrated GPU) driving the kiosk. Algorithm modules for video and audio analytics are deployed as docker containers in the existing CPU. These modules process:

- Video frames from cameras
- Audio frames from microphones

Real-time face, gesture, speech, and keyword recognition allow safe and secure interaction with the applications. Application actions can be performed comprehensively and easily. The **Authenticate > Navigate > Operate > Actuate** framework enables any application to become contactless. Refer to Figure 1.

A well-documented wrapper architecture converts all the AI events into application events. Application developers can quickly integrate these events with existing applications to deliver a brand-new experience to customers.

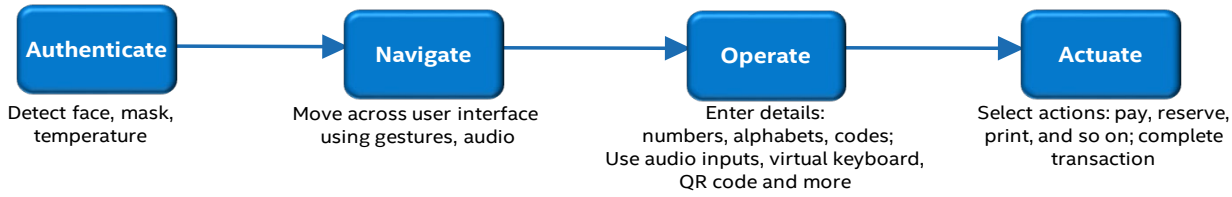

Figure 1: Operation flow of Outdu's contactless solution

User inputs—in the form of hand gestures, QR code on mobile phones, and speech—are analyzed and converted into navigation commands or keyboard inputs (numbers, alphabets). The analyzed inputs are placed in a message queue (called RabbitMQ).

The existing application running on the kiosk reads the message queue to receive user input. It uses a lookup table to convert new user inputs into existing commands. Finally, the application processes the user inputs as before with no change to existing application functionality and code. Table 1 sums up the events implemented by Outdu for their contactless solution

Table 1: Gesture, face, and audio events supported by Outdu's contactless implementation

GESTURE EVENTS	FACE EVENTS	AUDIO EVENTS
OK	Face detection	Voice detection: Initialization keyword (Marvin)
Palm up, down, left, right, push	Mask detection	Speech recognition: Numbers, Alphabets (English)
Two-finger click, up, down, left, right	Temperature check	Keyword detection: Up, Down, Left, Right, Yes, No, Off, Go, Next, Stop, Marvin
Thumb up, thumb down	Face recognition (optional)	
Two-hand zoom in, zoom out	Face count (optional)	
Cross hands, grab, fist		
Rotate hand clockwise, anti-clockwise		

Technologies Implemented

Custom all-in-one hardware developed using commercially available devices are connected to Intel® NUC to execute various workloads.

Outdu has integrated their deep neural networks using the Intel® OpenVINO™ toolkit to analyze video frames and audio singles. The neural networks are previously trained with public and custom datasets collected and curated by Outdu.

Figure 2 shows the hardware diagram for Outdu's contactless solution.

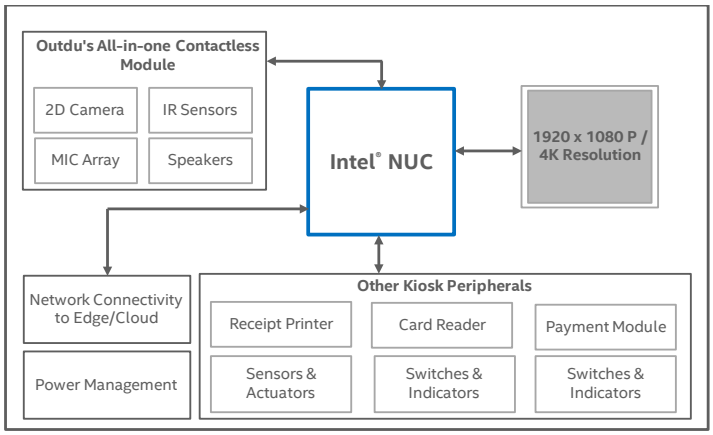

Figure 2: Intel® NUC and other HW in Outdu's contactless solution

While Outdu uses the Intel® NUC platform for their contactless development, other OEMs/SIs can use any Intel-powered platform for their implementation. Refer to ABOUT INTEL section for other recommended platforms.

The software models are based on Convolution Neural Network (CNN) architecture for image processing, Bidirectional Long Short-Term Memory (BLSTM) architecture to process hand gestures, and CNN and LSTM for speech recognition.

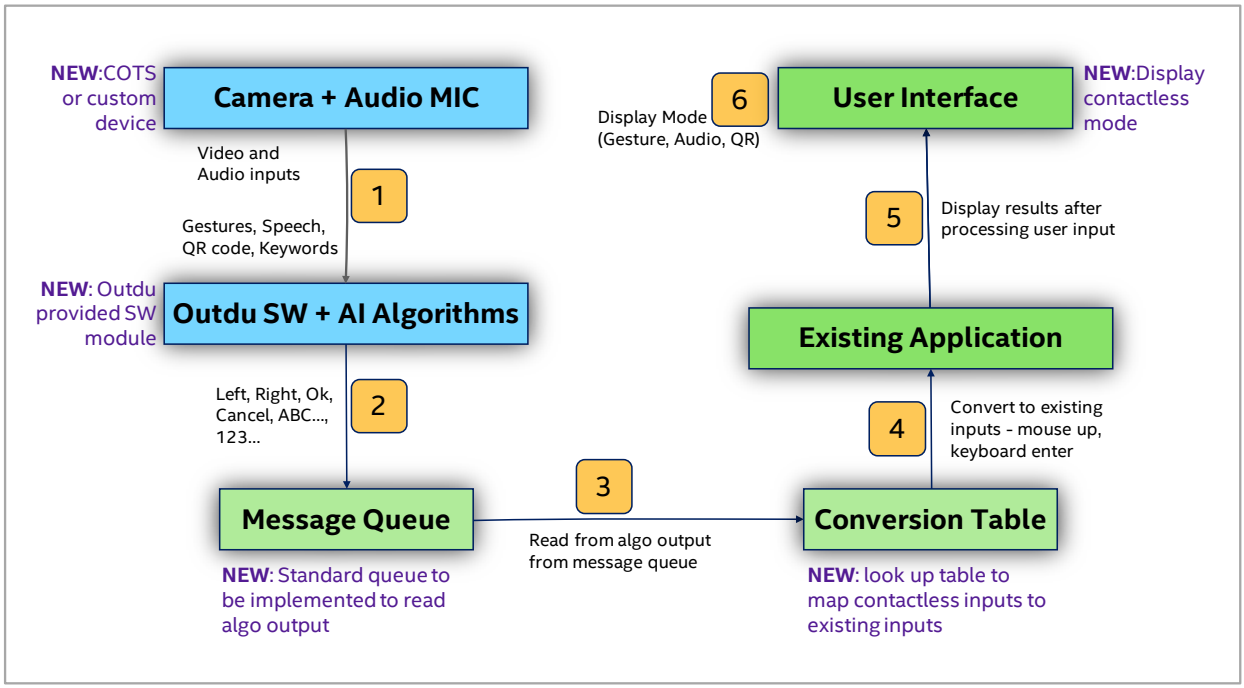

Figure 3: Application overview and execution flow

The solution components integrated using Intel® OpenVINO™ toolkit, to develop the AI pipeline using the camera and audio MIC, are shown in Figure 4. These models and SW orchestration support Linux* and Windows* operating systems.

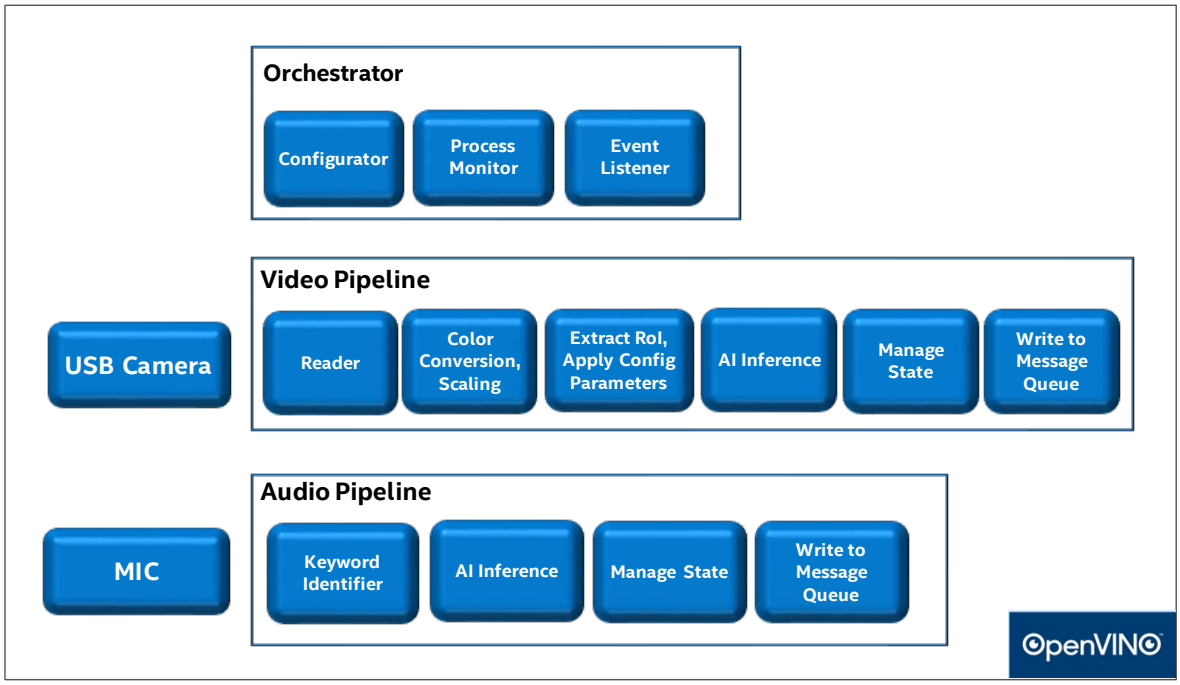

Figure 4: AI pipelines integrated using Intel® OpenVINO™ toolkit

The execution flow for processing audio and video inputs is illustrated in Figure 5.

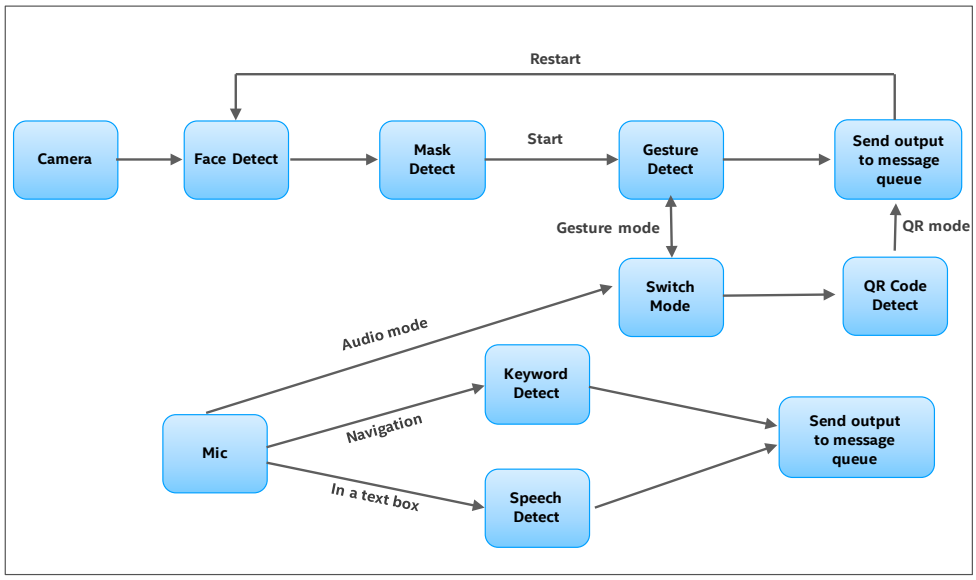

Figure 5: Execution flow of audio and video inputs

Benefits of Outdu's Contactless Solution

- **Contactless & Safe:** Users can interact with self-service terminals and kiosks without touching the screens for navigation and inputs.
- **Enforces Safe Operations:** User inputs are accepted only after a face is detected wearing a mask.
- **Intuitive Interactions:** Simple intuitive hand gestures are used to perform operations such as scrolling, moving, selecting, and canceling.
- **Virtual Keyboard:** For more specific inputs, a virtual keyboard on the screen can be operated with hand gestures to select specific keys, enter, cancel, and so on.
- **Speech Inputs:** Audio keywords are used to navigate, and speech is used to enter specific data in the form of alphabets and numbers.
- **Uses Existing Application:** No change to existing applications; the same screens, functions, and features can be carried forward.
- **Minimal Effort to Create Interface Application:** Create interface applications to read user inputs from a message queue and use a lookup table to convert contactless inputs into existing commands. These commands are processed by the existing application code as before.
- **Easy to Deploy:** Convert existing touch-based screens into touchless ones easily. First, add a USB camera and microphone to the existing hardware. Next, deploy AI analytics as a docker container to read from the message application.

Use Case Examples

Outdu's contactless solution is ideal for organizations looking to convert interactive touch-based kiosks into contactless ones at minimum cost and effort. Especially those organizations with high foot traffic having already deployed a large number of kiosks.

Examples of potentially convertible interactive kiosks can be found at:

- Airports and transit stations
- Restaurants
- Resorts and parks
- Large-format retail and brand stores
- Service centers
- Hotels
- Museums
- Malls
- Banks
- Financial service providers

Conclusion

Contactless solutions are the next natural transition to improve user experience in the new normal. Any system—existing or new—powered by Intel® Processors, can be transformed into a contactless system. Outdu's solution, developed with Intel® Distribution of OpenVINO™ toolkit, enables organizations to adapt to the contactless mode of operation for their existing devices. Moreover, Outdu's solution can be used for their planned deployment of self-service terminals and kiosks. Organizations can thereby continue to mobilize their existing investments in a safe and compliant manner and win the confidence of their users by using advanced technology solutions to minimize health risks.

ABOUT INTEL

You may know us for our processors. But we do so much more. Intel invents at the boundaries of technology to make amazing experiences possible for business and society, and for every person on Earth. By harnessing the ubiquity of the Internet of Things, the latest advances in memory and programmable solutions, Intel is disrupting the retail industry. Our rich portfolio of AI technologies is helping solve the toughest challenges.

Recommended Intel Compute Platforms for contactless kiosks/digital signages:

- Intel® Open Pluggable Specification
- Intel® Smart Display Module
- Intel® Smart Kiosk Module

ABOUT OUTDU

Over the past decade, Outdu has designed, deployed, and supported telco-grade video logistics, analytics, and reporting solutions for clients across multiple industries. Outdu offers AI/ Deep Neural Network-based video and audio analytics to support deployment of contactless applications. Outdu's offerings include AI algorithms, software products, and integrated HW and SW products. Visit outdu.com to learn more.

Other recommended Intel Compute Platforms:

- Intel® NUC
- Mini-ITX motherboards
- Third-party compute platforms based on Intel® Core™ processors

Intel Resources:

- Intel® Distribution of OpenVINO™ Toolkit
- Intel® OpenVINO™ Pre-trained Models
- Intel® Core™ Processors
- Intel® Vision Accelerator Design Products
- Intel® RealSense™ Technology
- Intel® RealSense™ SDK 2.0

OpenVINO™ toolkit

The Intel® Distribution of OpenVINO™ toolkit is a free, downloadable toolkit that helps developers fast-track the development of high-performance computer vision and deep learning into vision applications. It enables deep learning on hardware accelerators and streamlined heterogeneous execution across multiple types of Intel® platforms. It includes the Intel® Deep Learning Deployment Toolkit with a model optimizer and inference engine, along with optimized OpenCV* and OpenVX* libraries and functions for computer vision. This comprehensive toolkit supports the full range of vision solutions, speeding computer vision workloads, streamlining deep learning deployments, and enabling easy, heterogeneous execution across Intel platforms from device to cloud.