Case Study

Internet of Things
Healthcare

TheBlue.AI accelerates medical data anonymization in Nicklaus Children's Hospital

Partners TheBlue.AI, Intel, and apoQlar have collaborated on a solution for Nicklaus Children's Hospital in Miami, Florida that uses artificial intelligence (AI), mixed reality, and the cloud to help anonymize[1] personally identifiable information (PII).

> "It'll be revolutionary to be able to have people look at an operation in 3D. Then, when you think about the potential for surgeons to have a model they can operate on in their hands while watching, suddenly we have a learning environment that's unparalleled."
>
> **Dr. Robert L Hannan,**
> Director of the Cardiovascular Surgery Innovation Laboratory, Nicklaus Children's Health System

Healthcare organizations are looking to innovative technologies like mixed reality to plan operations and train doctors more effectively. However, many hospitals, medical facilities and pharmaceutical companies already struggle with handling their patients' personal information. Increasing pressure from regulation, new reporting standards and higher expectations of privacy from patients mean that these innovative technologies can cause compliance problems for hospitals.

Blue.GDPR is a solution designed to anonymize personal data in videos, images, medical data, documents, and speech. Based on machine learning algorithms, accelerated by Intel distribution of OpenVINO toolkit, it can help find and mask personally identifiable information (PII) like faces, names, and addresses in large volumes of structured and unstructured data.

Challenge

Mounting regulatory pressure means that healthcare organizations must find ways of handling PII without compromising the privacy of their patients.

- Reliable anonymization of data poses computational challenges for hospitals, many of which already have complex IT environments.
- Nicklaus Children's Hospital uses VSI HoloMedicine, a solution that enables surgeons to transform 2D MRI, CT, and other DICOM images into 3D holograms in real space using Microsoft HoloLens mixed reality glasses. However, to ensure video from the glasses can be stored for medical documentation, used as training material, and for analytics purposes, the data must be anonymized.

Solution

- Running on Intel® hardware on the Microsoft Azure cloud platform, Blue.GDPR obscures faces, redacts medical documents, and anonymizes patient data to make them unrecognizable. This allows videos and images to be processed without identifying individual members of staff or patients.

Results

- Nicklaus Children's Hospital is free to make use of VSI Holomedicine while planning operations and training doctors, while ensuring it is able to comply with privacy rules and regulations.
- The hospital is already planning an expansion of its training programs.

Medical data: Huge potential, but concerns over privacy

Medical data is of extremely high value: for healthcare organizations, it holds the potential of saving lives, educating staff and developing new treatments. But it is also incredibly sensitive - legal regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the US and the General Data Protection Regulation (GDPR) in the European Union impose strict rules regarding the handling of personal data. Contravening these regulations can result in heavy fines.

Compliance is also becoming harder as hospitals are increasingly forced to deal with third parties like insurance companies that require access to medical documentation, or patients requesting access to their own data. To add to this problem, the proliferation of new healthcare technology and practices are generating not only more data, but also new kinds – 3D cameras used for patient monitoring, for example.

The best way for hospitals to avoid regulatory issues is to help anonymize[1] data before it is shared by eliminating PII, but the process of anonymization of medical data poses challenges and has its limitations, especially when it comes to video data and genomic information.[1] It requires highly specialized AI models, since the datasets involved with medical data are unique (e.g. videos made during surgeries with visible inner organs). The models also need to be robust – if a model fails to anonymize even a few frames of a video, the privacy of the patient may be compromised.

Additionally, new IT solutions typically become part of existing complex IT environments in hospitals. Factors such as internal regulations connected with software standards, and a lack of hardware specifically optimized for running AI algorithms, means that reliable anonymization is challenging from a computational perspective.

Nicklaus Children's Hospital: a center for innovation

Nicklaus Children's Hospital in Miami is South Florida's only licensed specialty hospital exclusively for children, with nearly 800 attending physicians and more than 475 pediatric subspecialists. As a highly innovative hospital - home to the largest pediatric teaching program in the southeastern United States - it applies many high-tech solutions to support operations and facilitate training.

One such solution came when the hospital decided to implement VSI HoloMedicine, a solution built by Hamburg-based firm apoQlar. It allows physicians to use 3D images of patients, viewed through mixed reality glasses superimposed on the actual patient, to visualize, highlight, and measure anatomical structures. It has a variety of use cases and is particularly helpful when planning operations and for training new doctors.

To help guarantee the proper handling of personal information in the new types of data created, VSI Holomedicine makes use of Blue.GDPR, a solution built by TheBlue.AI that can reliably anonymize[1] personal data in videos, images, medical data, documents, and audio. This ensures the data can be used as reference and training material, as well as be the subject of further analytics scenarios. Using an AI algorithm, Blue.GDPR obscures faces to make them unrecognizable, allowing videos and images to be shared with minimal risk to staff or patient data privacy.

Normally, AI models are trained to recognize faces through standard ratios and features common on all human faces. In an operating room, where doctors and nurses are wearing protective equipment, standard facial features are no longer visible. To better cater to the medical field, the algorithm behind Blue.GDPR was extensively trained with data and images specific to the medical setting, allowing it to optimally identify and obscure all faces.

Technical solution details

This section reviews the hardware, software and optimization frameworks that make this solution possible. It is designed to help software developers and solution builders who are looking to deploy similar programs in their own organizations.

Building a dataset
TheBlue.AI constructed a dataset consisting of frames derived from short video recordings, each depicting a scene involving a real medical scenario like surgery, medical exams, or training. To increase reliability, videos were recorded in several different medical institutions. All videos were recorded with Microsoft HoloLens devices using the VSI Holomedicine application.

Frames were labeled by hand, designating each human head in frame, as well as the shape of each face. The resulting dataset consisted of 1613 individual images, with around 3500 objects labelled.

Topology
TheBlue.AI used a variant of a popular and powerful topology for instance segmentation – Mask R-CNN. Mask R-CNN is a two-stage detector with an additional instance segmentation component. The first stage, called region proposal network (RPN) is responsible for generating candidate bounding boxes for objects to be detected. The second stage – RoIPooling – predicts, classes, and refines the locations of those boxes.

The input to the detector comes from a Resnet-50 convolutional neural network (CNN) enriched by an additional module – the Feature Pyramid Network (FPN). FPN creates pyramid features build on top of intermediate feature maps coming from each stage of Resnet. This increases the overall inference speed and accuracy of the model.

Then, TheBlue.AI employs a variant of Mask-RCNN implemented in Intel® Distribution of OpenVINO™ Training Extensions toolkit, that is built on top of a Pytorch framework. The main difference from the original version is the number of feature maps processed in FPN, RPN and RoiHeads blocks, which is decreased by the factor of two. This results in amazing, fast processing times with minimal loss of accuracy.

Figure 1. Input (left) and anonymized output (right) of Mask-RCNN.

Intel® Distribution of OpenVINO™ Toolkit

The solution uses CPU-based Azure instances equipped with Intel® technology, including Intel® Core™ i7 and i5 processors. The Azure platform enables developers to take full advantage of the Intel Distribution of OpenVINO Toolkit. Based on convolutional neural networks (CNN), the toolkit extends processing across Intel® hardware to boost performance for demanding workloads. By using the toolkit, TheBlue.AI was able to output extremely high frames-per-second (FPS) of anonymized[1] data.

Optimization and Testing

The original model was trained on Pytorch and serves as a benchmark for subsequent tests. The Pytorch model was exported to ONNX format and then to OpenVINO Intermediate Representation (IR). Intel® Distribution of OpenVINO™ Toolkit version 2020.4.287 was used. To ensure reproducibility all inference tests were conducted using the Intel® DevCloud for the Edge.

Tests conducted in pure Pytorch followed those contained in benchmark script: batch size was 1, iterations were 10 and the FPS was expressed as batch size * iterations/total duration. Input image size was 480x480 pixels.

> "It's the ultimate form of visualization. Every sports star visualizes kicking the winning goal – well now he can visualize kicking the winning goal wearing a Hololens."
>
> **Dr. Robert L Hannan,** Director of the Cardiovascular Surgery Innovation Laboratory, Nicklaus Children's Health System

Looking forward

Nicklaus Children's Hospital can continue pioneering the use of mixed reality for medicine, and thanks to Blue.GDPR is able to help do so with a reduced risk of regulatory issues. It now plans to make 3D videos of open-heart operations available online for training.

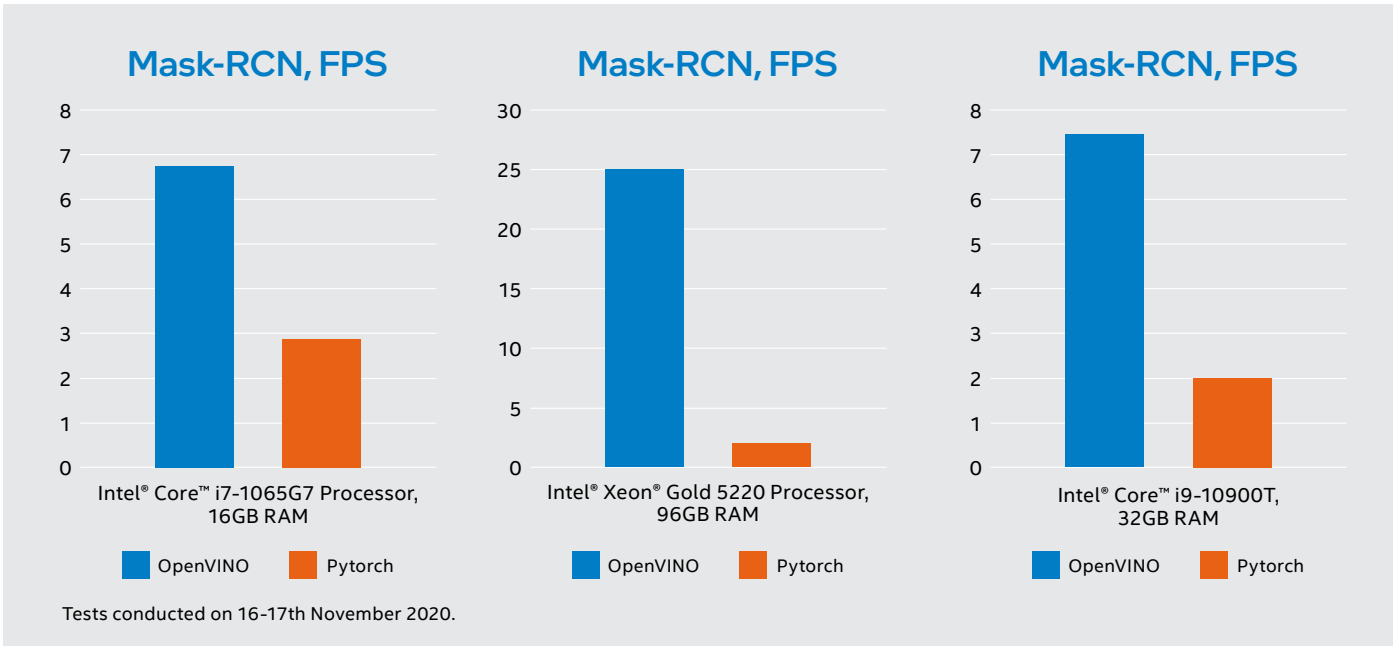

Figure 2. Testing data on three commonly used business-class and server CPUs.

About TheBlue.AI

TheBlue.AI is an AI technology provider based in Hamburg. Their Blue.GDPR technology allows the anonymization of personal and industry specific information within documents, images, videos, and audio data.

About apoQlar

apoQlar GmbH is a healthcare technology provider, specializing in mixed and augmented reality. Their solution VSI HoloMedicine allows surgeons to see anatomical structures during surgical planning - helping them to improve precision, save time, and avoid post-operation procedures.

About Microsoft Azure

Microsoft Azure is a public cloud computing platform. Intel and Microsoft collaborate closely to ensure that the platform is optimized for the latest Intel® technologies, delivering a broad portfolio of intelligent services.

Learn More

You may find the following resources helpful:

- **Intel Distribution of OpenVINO Toolkit**
- **Intel DevCloud for the Edge**
- **Blue.GDPR**

[1] Genomic Data cannot be anonymized.

0821/EB/CAT/PDF ♻ Please Recycle 347911-001EN