Save Lives with Machine Vision and Audition

These projects use Machine Vision and Audition techniques to improve patient care during EMS emergency responses.

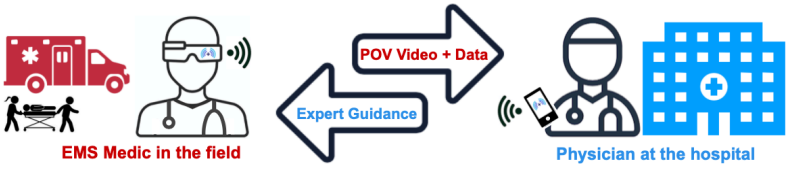
You will use the audio/video stream from a medic's headcam to:

- Look for specified medical device displays and extract the information on them;
- Listen to verbal interactions among the medics and patients to identify and extract key information;
- Look for driver's license and other IDs to extract key information for the medical record

1

Overview: Saving Lives with Machine Vision & Audition

JHU docs have developed a **smart glasses-based platform** that lets physicians in the hospital provide **support and guidance for EMS teams** when responding to medical emergencies in the field.



The remote physicians can see and hear what the medic sees and hears, and that means **faster diagnosis and start of treatment**.

Physicians can also **guide critical procedures** like intubations, which have a low success rate outside the operating room.

2

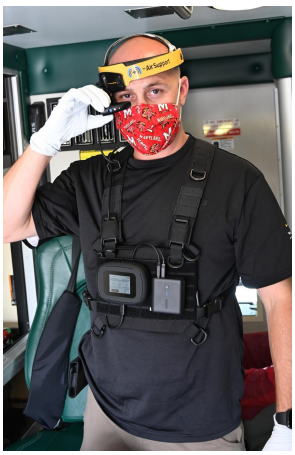
Overview: Three Challenges

(1) To further improve care, the remote physician should have direct access to the **information available from medical devices** used by the medics, especially vitals signs monitors and ultrasound scanners. Unfortunately there are poor interoperability standards for medical devices. But the **human-readable displays of these devices** will be in view of the smart glasses camera. ***Can machine vision recognize those displays and ingest those data?***

(2) Streamlining patient identification by pulling key information from standard documents like driver's licenses means that medics spend less time on paperwork and more time treating patients. ***Can machine vision recognize these documents and comprehend and ingest their contents?***

(3) To ensure continuity of care, identifying and documenting **key verbal interactions** of the medics and patients is crucial. Recognizing and documenting **non-verbal respiratory symptoms** like coughing is also important. ***Can machine audition and NLP do this?***

Per the notes in my email, I think we're not ready for (3) and I will pull this from the presentation unless I hear otherwise.



Dr. Nick Dalesio of JHU in an ambulance, wearing smart glasses on a medical head mount. On his chest are a MiFi unit for mobile broadband and an auxiliary battery.

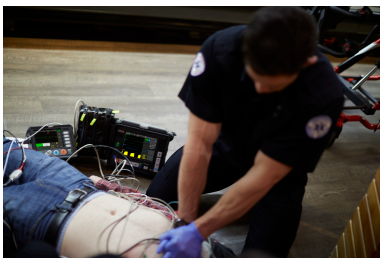
3

1. Medical Devices – Recognize Visual Display and Ingest Data

Machine Vision, Object Recognition & Tracking, OCR

Goals: Ingest data from the visual displays on medical devices used for patient care.

- Recognize the display of a vitals signs monitor and an ultrasound device from video streaming from of a wearable camera on a medic who is in motion while treating a patient
- Capture relevant image frames/video when a display monitor is on camera so:
  - Alpha-numeric vitals signs data can be "read" and translated to telemetry.
  - Ultrasound images/video are captured and any alpha-numeric data is translated to telemetry
- Include metadata required to sync back to the original video(s) from the camera(s), regardless of the kind of output (telemetry, image, video, etc.)



4

1. Medical Devices – Recognize Visual Display and Ingest Data

What You Will Do

- Create of an algorithm that will recognize and track the human-readable visual displays of portable ultrasound and vitals signs monitors* as ”seen” in video from a head-mounted camera.

Deliverables

- Machine vision algorithm capable of recognizing, tracking, and ingesting information from the human-readable screens of ultrasound and vitals signs monitors from a video stream.
- A method for making the ingested data available to another system.

Group Size

- 1-2

Skills

- Machine Vision: Object Recognition & Tracking, OCR

Mentors

- Nick Dalesio, MD, MPH nick.dalesio@mdairsupport.com; Laeben Lester, MD laeben.lester@mdairsupport.com

** If focus on only one type of device is appropriate for the scope of the project, ultrasound is the priority.*

5

2. IDs – Recognize and Extract Key Data

Machine Vision, Object Recognition, OCR

Goals: Streamline the process of recording and inserting information into the medical note (the record of the encounter) based on specified patient and EMS identification documents.

Specified IDs will typically be driver’s licenses of patients and EMS IDs of medics. However, other forms of government ID like Tribal/Native American ID cards, voter ID cards, school IDs and others can be expected.

- Recognize a specified ID type when it is held up to the camera
- Automatically extract information about the person
  - Recognize categories of information and store them accordingly
    - e.g. understand what is the name, address, age, donor status, etc.
    - Recognize and apture the person’s photo as an image and, separately, entire license as an image
    - Recognize, capture and separately store images of any BAR or QR codes on front

6

2. IDs – Recognize and Extract Key Data

What You Will Do

- Create an algorithm that can recognize a driver's license when it is held up to view in a head-mounted video camera, understand and extract visual and textual information from the drivers license.*

Deliverables

- An algorithm that will automatically capture all information from the front of a driver's license when it appears in a video stream, to recognize, extract, categorize and store all visual and textual information (e.g., full image of license, person's photo, BAR/QR codes, name, date of birth, address, etc.)

Group Size

- 1-2

Skills

- Machine Vision: Object Recognition & Tracking, OCR

Mentors

- Nick Dalesio, MD, MPH nick.dalesio@mdairsupport.com; Laeben Lester, MD laeben.lester@mdairsupport.com

* If the scope of the project permits, extend this capability to other forms of identification including EMS ID's, other medical ID's, student ID's or any other form of identification available to the group.

Per the notes in my email, I think we're not ready for (3) and I will pull this from the presentation unless I hear otherwise.

3. Voice – Recognize and Track Key Interactions

Machine Audition, Voice Recognition, Speech-to-text, Natural Language Processing

Goals: Streamline the process of recording and inserting information into the medical note based on verbal interactions with the patient.

- From a recording of a medical encounter, "listen" to voices and "follow" conversations to identify and note:
  - Medic's question and patient's/other's answer to specific and common questions from the medic. For example, *Any allergies to medicine? What's your date of birth? Do you have a pacemaker?*
  - Less structured interactions, like a patient's description of *how they got injured, their current condition, their stated preferences for care, level of pain,* etc.
  - Any patient coughing, wheezing (by isolating this, it can be analyzed later for diagnosis)
  - Medic's description of care. For example, the medic may say things like, *I'm going to give you an oxygen mask now. I'm going to put a neck brace on you. We need to intubate. Can you move your toes? Get the AED. We're taking you to the emergency room now.*
- All of the above to be made available:
  - In transcript form, labelling unique voices consistently throughout (voice 1, voice 2, etc.)
    - With an attempt to identify the unique voice's role, e.g., medic 1, medic 2, patient, etc. or even identity, e.g., "Medic Jane Doe" or "Patient John Doe" etc.
    - Also flagging discrete interactions within the transcript, e.g., a question asked by a medic and the corresponding answer from a patient.
  - for the transcript and start/stop times for discrete interactions With any relevant metadata including timecode so that other systems have what they need to:
    - Find the relevant start/stop point of a discrete interaction in the original audio/video source
    - Extract the corresponding audio/video clip for review or insertion into a medical record.

Per the notes in my email, I think we're not ready for (3) and I will pull this from the presentation unless I hear otherwise.

3. Voice – Recognize and Track Key Interactions

Machine Audition, Voice Recognition, Speech-to-text, Natural Language Processing

I will complete this section if this project (3) makes the cut for this semester.

What You Will Do

- x

Deliverables

- x

Group Size

- 2-6

Skills

- Machine Audition: Voice Recognition, Speech-to-text, NLP

Mentors

- Nick Dalesio, MD, MPH nick.dalesio@mdairsupport.com; Laeben Lester, MD laeben.lester@mdairsupport.com

9

One More Thing…

The projects described here will help build the foundation for our next phase, which is to use these and other inputs to provide **AI-based diagnostic and decision support** for medics and nurses, in addition to the remote video support from human physicians we enable today.

We've got a lot to do and we hope you will consider joining us on our mission to improve healthcare outside of the hospital, especially for people in rural and underserved communities who rely on under-resourced clinics, nursing homes and other facilities for their care.

10