

AWARD NUMBER: W81XWH-17-C-0238

TITLE: Complete and Resilient Documentation (CARD) for Operational Medical Environments

PRINCIPAL INVESTIGATOR: Kuang-Ching Wang, PhD

CONTRACTING ORGANIZATION: Clemson University, Clemson, SC

REPORT DATE: December 2021

TYPE OF REPORT: Final Report

**PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012**

**DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited**

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				<i>Form Approved</i> OMB No. 0704-0188	
1. REPORT DATE December 2021		2. REPORT TYPE Final		3. DATES COVERED 25Sep2017 - 30Sep2021	
4. TITLE AND SUBTITLE Complete and Resilient Documentation (CARD) for Operational Medical Environments				5a. CONTRACT NUMBER W81XWH-17-C-0238	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Kuang-Ching Wang, PI kwang@clemson.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Clemson University Sponsored Programs 201 Sikes Hall Clemson, SC 29634-0001				8. PERFORMING ORGANIZATION REPORT	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release, Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT In response to the HITI HFEHRI solicitation for research to demonstrate and validate hands-free electronic health record data entry solutions that will operate reliably in noisy operational environments, alleviate disruption of care for documentation, and prevent loss of documentation, this project studied a system-oriented approach to meet these objectives with a platform, CARD, aimed to enable resilient hands-free data collection, preserve complete documentation through stages of care, and present timely information useful for the medical operation. The wearable CARD platform was developed with a GoPro camera, binaural microphones, an NVIDIA TX2 embedded GPU computer, SDXC memory, and Wi-Fi network interface, running Kaldi-based speech recognition with deep neural networks pre-trained on a powerful NVIDIA DGX 8-GPU computer. The entire system can be viewed and controlled by a commander dashboard. The team developed and published several speech enhancement techniques that significantly improved the speech recognition performance, while it continued to develop machine learning algorithms for post-recognition error detection and correction.					
15. SUBJECT TERMS Battlefield, Hands-free, Electronic health record, Documentation, Operational environments, Medical Workflow, Stages of care, Speech recognition, Noise, GPU, Deep learning, Neural network, Video, Simulation					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	98	19b. TELEPHONE NUMBER (include area code)

TABLE OF CONTENTS

	<u>Page</u>
1. Introduction	4
2. Keywords	4
3. Accomplishments	4
4. Impact	14
5. Changes/Problems	16
6. Products	16
7. Participants & Other Collaborating Organizations	18
8. Special Reporting Requirements	22
9. Appendices	23

1. INTRODUCTION:

In response to the HFEHRI (Hands-Free Electronic Health Record Data Entry Initiative) solicitation for research to demonstrate and validate hands-free electronic health record data entry solutions that will operate reliably in noisy operational environments, alleviate disruption of care for documentation, and prevent loss of documentation, this project studies a system-oriented approach to meet these objectives with a platform aimed to enable resilient hands-free data collection, preserve complete documentation through stages of care, and present timely information useful for the medical operation.

2. KEYWORDS:

Battlefield
Hands-free
Electronic health record
Documentation
Operational environments
Medical
Workflow
Stages of care
Speech recognition
Noise
GPU – Graphics processing unit
Deep learning
Neural network
Natural language processing
Video analysis
Simulation

3. ACCOMPLISHMENTS:

What were the major goals of the project?

As originally conceived in August 2016, the project's goal was to develop a platform that would achieve three objectives - **enabling** resilient hands-free data collection, **preserving** complete documentation through stages of care, and **presenting** timely information useful for the medical operation using deep learning and natural language processing techniques. To achieve our goal, the **CARD technology platform** was proposed, consisting of GoPro cameras, GPU-based embedded computer, onboard non-volatile data storage and wireless networks, to accomplish the three objectives. To assure our solution would be anchored to **clinical and operational requirements**, a collaborative team bringing together *Clemson University researchers* and *Palmetto Health emergency medicine physicians and nurses* was composed to execute the project following four specific aims (as defined in the Statement of Work):

- **Medical Workflow Studies** – Using focus group interviews, EMS simulation drills, and medical-workflow-driven performance metrics for CARD, this aim informs the design of the CARD platform and assessment methods.
- **Platform Development** – Using the integrated CARD platform with deep neural network (DNN) based learning based automatic speech recognition (ASR) and video analysis software trained for military operation conditions and DoD's digital medical data platforms (TCCC card and Cerner EHR), this aim creates an extensible technology prototype to demonstrate the complete CARD workflow.

- **Assessment** – Using medical workflow performance metrics and the CARD platform, this aim assesses performance in the data documentation, speech and video analysis outcomes, and the level of acceptance by EMS professionals in simulated medical operations.
- **Roadmap Development** – This aims to derive recommendations of a technology maturation roadmap for the CARD technologies.

What was accomplished under these goals?

The project ran from September 2017 to September 2021 (Original 2-year project duration with two 6-month no-cost extensions and one 1-year no-cost extension). The medical workflow and the CARD platform (Aim 1 and 2) have been completed within the first two years, while the two extension years have been devoted to further advancement of the speech recognition technology with deep neural network and natural language processing methods. Figure 1 and 2 show the CARD platform architecture, hardware, a real-time operation dashboard, and the speech recognition output display. The wearable CARD platform was developed with a GoPro camera, binaural microphones, an NVIDIA TX2 embedded GPU computer, SDXC memory, and Wi-Fi network interface, running Kaldi-based speech recognition with deep neural networks pre-trained on an NVIDIA DGX 8-GPU computer.

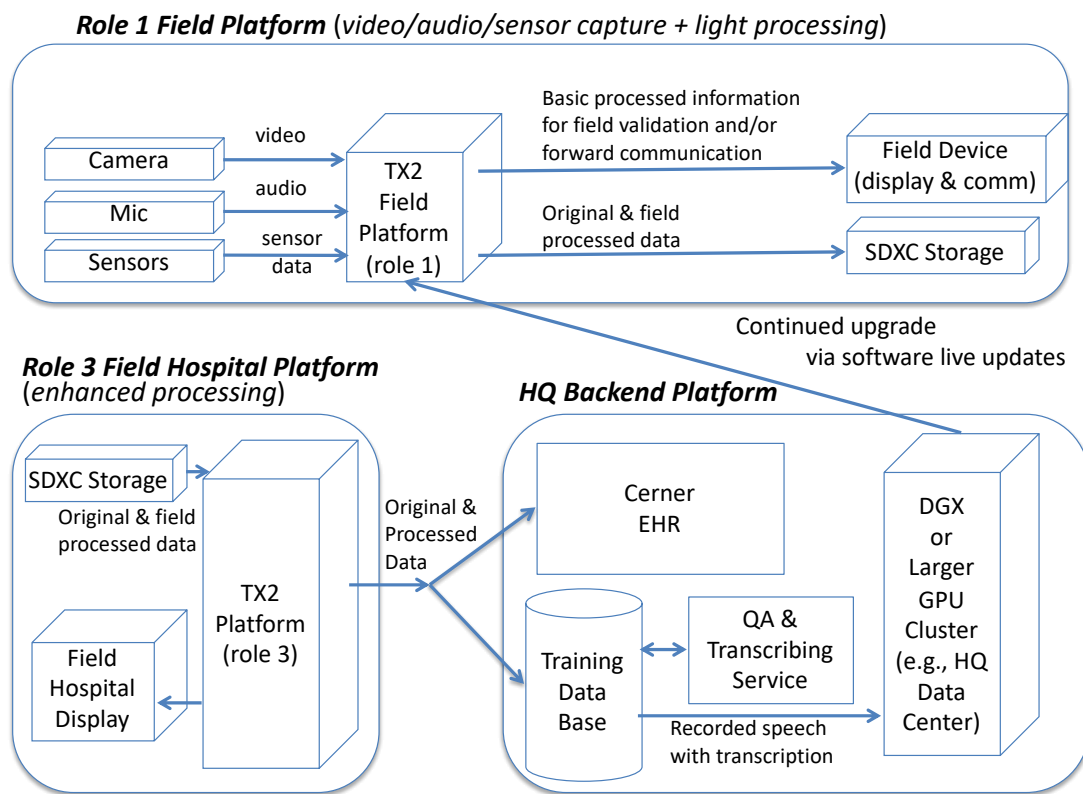
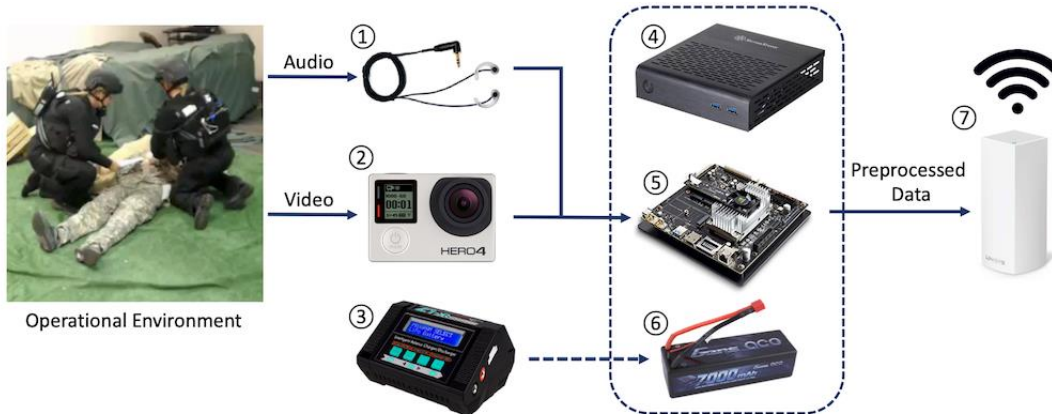


Figure 1. CARD platform architecture



Device	Product	Function
①	SP-TFB-2 Sound Professionals In-ear Binaural Microphones	Audio streaming capture for automatic speech recognition
②	GoPro HERO4 v05.00	Video streaming capture & real-time monitoring
③	EV-Peak C1-XR Charger	LiPo battery charger for field platform
④	SilverStone ITX Case PT13B-120	Enclosure for the field platform
⑤	NVIDIA Jetson TX2	Preprocessing of captured data
⑥	Gens LiPo Battery Pack 7000mAh 14.8V	Continuous power supply up to 6 hours
⑦	Linksys Velop WHW0303 Router	Secure network connection between field and field hospital

(a)

Device	Runtime	Memory	Decoded	Recorded	TCCCProgress
CAM1	00:43:54	74%	00:17:39	00:25:33	337 / 336 Recorded
CAM2	02:32:15	21%	01:14:34	01:52:28	67 / 336 Recorded
CAM3	02:24:48	20%	01:00:36	01:44:39	67 / 336 Recorded
CAM4	00:18:32	82%	00:04:23	00:06:36	12 / 336 Recorded

Below the table are camera feeds for CAM1, CAM2, and CAM3, and patient identification labels for CAM1, CAM2, and CAM3.

(b)

Device ID	Text Output
CAM 1	cut on oh hey hey hey what's your name we got bleeding so like we need to turn it <unk> i <unk> uh for the store to get on okay alright alright he's <unk> he's got good radio closest <unk> you know where you are right now oh okay <unk> okay <unk> the blood blood <unk> friends last time you a <unk> or
CAM 2	get ready to go to the helicopters cats have already been called i said we could probably get ready to move now you got all the history we need one [laughter] casual collect any allergies so i kinda do that but i ain't got turn around right [noise] right <unk> [laughter] that'll last their number
CAM 3	like she has a heart rate and the one thirties probable radio pulses by letter like kids alert <unk> no evidence of head trauma allergic to penicillin and takes out are all pretty she knew whether medical
CAM 4	good so we're going to have to start getting ready to move well okay thank you that he didn't have any pulls down here obviously yep right good fleeting stops i'm gonna go ahead why ask more questions and go ahead and finish dressing as women right

(c)

Figure 2. (a) CARD platform, (b) CARD dashboard, (c) Speech recognition output display.

Throughout the project, 4 refereed conference papers have been published at highly ranked international conferences on our research findings for **DNN-based speech enhancement** and **systematic evaluation of ASR in medical environments**.

1. Ju Lin, Sufeng Niu, Zice Wei, Xiang Lan, Adriaan J Wijngaarden, Melissa C Smith, Kuang-Ching Wang, “Speech enhancement using forked generative adversarial networks with spectral subtraction,” in Proceedings of Interspeech 2019, 2019.
2. Ju Lin, Sufeng Niu, Adriaan J van Wijngaarden, Jerome L McClendon, Melissa C Smith, Kuang-Ching Wang, “Improved Speech Enhancement using a Time-Domain GAN with Mask Learning,” in Proceedings of Interspeech, 2020.
3. J Lin, AJ van Wijngaarden, KC Wang, MC Smith, “Speaker-aware Speech Enhancement with Self-attention,” in Proceedings of EUSIPCO (European Signal Processing Conference), 2021.
4. Snigdhaswin Kar, Prabodh Mishra, Ju Lin, Min-Jae Woo, Nicholas Deas, Caleb Linduff, Sufeng Niu, Yuzhe Yang, Jerome McClendon, D. Hudson Smith, Melissa Smith, Ronald Gimbel and Kuang-Ching Wang, “Systematic Evaluation and Enhancement of Speech Recognition in Operational Medical Environments,” in Proceedings of IJCNN (International Joint Conference on Neural Networks), 2021.

In addition, two refereed journal papers have been accepted (one with minor revision) in reputable medical and engineering journals on our research findings:

5. MinJae Woo, Prabodh Mishra, Ju Lin, Snigdhaswin Kar, Nicholas Deas, Caleb Linduff, Sufeng Niu, Yuzhe Yang, Jerome McClendon, D Hudson Smith, Stephen L Shelton, Christopher E Gainey, William C Gerard, Melissa C Smith, Sarah F Griffin, Ronald W Gimbel, Kuang-Ching Wang, “Complete and Resilient Documentation for Operational Medical Environments Leveraging Mobile Hands-free Technology in a Systems Approach: Experimental Study.” JMIR mHealth and uHealth, 9(10):e32301, 2021, doi: 10.2196/32301.
6. Ju Lin, Adriaan J. van Wijngaarden, Kuang-Ching Wang, and Melissa C. Smith, “Speech Enhancement Using Multi-Stage Self-Attentive Temporal Convolutional Networks.” IEEE Transactions on Audio, Speech and Language Processing. *Accepted with Minor Revision.*

The project has also presented both as a poster and an oral presentation at the 2017 and 2018 Military Health System Research Symposiums.

Research Accomplishment Details

Medical Workflow Studies

Focus group interviews of Emergency Medicine Services (EMS) members, transport nurses, and Emergency Department (ED) physicians (Figure 3) were conducted early in the project, confirming strong interests of the CARD platform design with the following highlighted recommendations:

- Verbal and non-verbal communication are both important. Often, many people would be talking, sometimes including multiple patients.
- Strong desire for sending advance patient status to physicians.
- Strong opinions on camera mounting positions and importance of on/off control.

	TOTAL n	Experience <10 years	Experience >10 years
EMS	6	17%	83%
Transport RNs	11	0%	100%
ED MDs	9	33%	67%
TOTAL	26		

Figure 3. Focus group interview demographics.

Three simulation drills with wounded patients (actors) and EMS team (real EMS staffs + ED doctors) were conducted in the University of South Carolina Medical Simulation Center throughout the project. Figure 4 shows photos of the EMS staff, helicopter transport nurses, and the ED doctors. The most important outcome from the drills is a total of 6+ hours of video and audio recordings of their rescue operations complete with synthesized battlefield noise in the background, scripted medic and patient conversations for three distinct trauma types: below-knee amputation with arterial bleeding secondary to an improvised explosive device injury; gunshot wound through and through chest with respiratory distress; blast injury with concussion (altered mental status) and tympanic membrane perforations. The data has been a valuable reference of the speech and video conditions expected in battlefield medical operations.



Figure 4. EMS drills conducted in the medical simulation center.

The “drill data” set has been used as the final test for the CARD platform performance. Table I shows CARD’s speech recognition performance using (a) the publicly available “Common Voice” data set versus using (b) the drill data. As can be shown, CARD performance against the drill data has a clear gap from the synthesized data with the most intense noise level (0 dB). Note that the observed 33% error rate with the drill data is comparable to that reported in other recent academic and industry speech recognition papers^{1,2}.

¹ Mani, Anirudh, et al. "ASR error correction and domain adaptation using machine translation." ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020.

² C. Liu, F. Zhang, D. Le, S. Kim, Y. Saraf and G. Zweig, "Improving RNN Transducer Based ASR with Auxiliary Tasks," 2021 IEEE Spoken Language Technology Workshop (SLT), 2021, pp. 172-179, doi: 10.1109/SLT48900.2021.9383548.

Model	0 dB	5 dB	10 dB	15 dB	20 dB
Intermittent					
Baseline	33.36	22.81	14.18	8.98	6.55
+ MTR	17.44	10.85	7.68	5.93	5.28
Continuous					
Baseline	77.61	51.43	28.15	15.83	9.51
+ MTR	30.69	17.18	10.18	7.47	6.31

(a)

Model	Original	S-ForkGAN	SEGAN
Baseline	69.85 (0.61)	57.04 (0.67)	62.37 (0.62)
+ MTR	43.04 (0.72)	51.88 (0.70)	57.99 (0.66)
+ LM	33.34 (0.81)	46.33 (0.78)	53.03 (0.75)

(b)

Table I. CARD speech recognition performance when tested with (a) the Common Voice data set with *intermittent noise* and *continuous noise* of various intensities (dB) versus (b) the drill data set. “Model” refers to the different neural networks using different training methods (see paper #4, Kar 2021 for details). *Baseline*: Trained with clean audio. *MTR*: Multi-style TRaining, i.e., trained with synthesized noise. *LM*: Language model enhancement. WER: Word Error Rate. Medical WER: Word Error Rate considering only medically relevant words.

Note that the Medical WER metric has been used in I(b). The decision to define this new metric has been influenced by CARD’s decision to integrate its speech recognized outputs with the DD Form 1380, TCCC Card as shown in Figure 5. The Medical WER assesses only medically related words that appear in the recognized texts.

The screenshot shows the TCCC CARD dashboard for patient James Tiger. The patient's name is James Tiger, and the battle roster number is 101. The patient is identified as 'Unidentified Patient' with the call sign 'CAM'. The injury is described as a right leg blast amputation. The dashboard includes sections for treatments, medications, and vital signs.

Fluid	Name	Volume	Route	Time
Fluid				
Blood Product				

MEDS:	Name	Dose	Route	Time
Analgesic (e.g., Ketamine, Fentanyl, Meperidine)				
Antibiotic (e.g., Moxifloxacin, Ertapenem)				
Other (e.g., TXA)				

Signs & Symptoms:	Time
Pulse (Rate & Location)	130
Blood Pressure	/ / /
Respiratory Rate	20
Pulse Ox % O2 Sat	
AVPU	
Pain Scale (0-10)	

Figure 5. TCCC card as displayed in the CARD real-time dashboard.

Platform Development & Assessment

Given the prevalence of speech recognition technologies in the industry, yet none of them were committing to support the military's operational use, we made the decision to focus on rigorous, systematic evaluation of such state-of-the-art solutions with the intent to identify key technical bottlenecks that the military can focus on to enable such solutions in the foreseeable future. Our work have therefore been conducted for two purposes: 1) systematic comparison of the state-of-the-art industry software "DeepSpeech" and open source software "Kaldi" performance with realistic battlefield medical operation audio recordings; 2) develop and demonstrate tangible enhancement methods and identify their critical enablers;

1. Performance comparison with industry's state-of-the-art speech recognition solutions.

We designed a systematic experimental study for investigating how military noise impacted a DNN-based ASR. We also evaluated how the addition of battlefield noise to the training set using multi-style training³ increased the robustness of the ASR model to noise. The performance improvement of the system by language model adaptation for the medical domain was also analyzed in this work. The overall performance of the hybrid Kaldi model⁴ was compared with the Mozilla DeepSpeech end-to-end model⁵ using WER as a standard metric. We also assessed experimentally how speech enhancement of the noisy data impacted ASR results.

The complete evaluation results were published in paper #4 listed above (Kar 2021). The performance of DeepSpeech and Kaldi models was tested in three experimental settings: 1) clean speech, 2) noisy speech at randomized SNR levels, and 3) noisy speech at specific SNR levels. The experimental results indicated that the proposed Kaldi model outperformed DeepSpeech when tested against both Switchboard and Common Voice datasets as shown in Table II. In Table III, when tested with the Common Voice dataset mixed with intermittent and continuous noise from 0 dB to 20 dB, Kaldi performed better than DeepSpeech. Table IV compares the WER performance under fixed SNR levels with the Common Voice dataset. An increase in accuracy is observed as the SNR level increases. Kaldi again gave better performance as compared to DeepSpeech. It is worthwhile to note that DeepSpeech is an "end-to-end deep neural network" based solution which is known to yield better performance when trained with "extremely large" data sets with "matching" audio properties for the targeted operating environment. **With smaller training data sets, we concluded that Kaldi consistently achieved better performance than DeepSpeech.**

³ R. Lippmann, E. Martin and D. Paul, "Multi-style training for robust isolated-word speech recognition", ICASSP'87. IEEE International Conference on Acoustics Speech and Signal Processing, vol. 12, pp. 705-708, 1987.

⁴ D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, et al., "The Kaldi Speech Recognition Toolkit", IEEE 2011 Workshop on Automatic Speech Recognition and Understanding. IEEE Signal Processing Society, December 2011.

⁵ A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, et al., "Deep speech: Scaling up end-to-end speech recognition", arXiv preprint, 2014.

Model	Common Voice	Switchboard
DeepSpeech	13.40	29.01
Kaldi	5.14	9.04

Table II. WER (%) for DeepSpeech and Kaldi with clean speech.

Model	Intermittent Noise	Continuous Noise
DeepSpeech	22.03	33.01
Kaldi	11.39	18.92

Table III. WER (%) for DeepSpeech and Kaldi with randomized noise intensities.

Model	0 dB	5 dB	10 dB	15 dB	20 dB
<i>Intermittent</i>					
DeepSpeech	27.55	17.33	18.17	16.90	15.66
Kaldi	17.44	10.85	7.68	5.93	5.28
<i>Continuous</i>					
DeepSpeech	44.09	29.72	21.83	17.94	15.98
Kaldi	30.69	17.18	10.18	7.47	6.31

Table IV. WER (%) for DeepSpeech and Kaldi with specified noise intensities.

2. Speech enhancement

Speech enhancement refers to the suppression of noise in audio signals prior to applying it to the speech recognition model. Given the significance of noise impacts in the military environment, this turned out to be the most fruitful area of research in the CARD project and we achieved significant success that captured attention at a series of international conferences and journal publications (papers #1-3, #6 above).

First, we investigated the general adversarial network (GAN) based speech enhancement. GAN consists of two neural networks competing with each other in a two-player min-max game. We proposed an enhancement over the speech enhancement generative adversarial network (SEGAN) model published in 2017⁶.

Next, we discovered in our drills that the speech recognition performance deteriorated when multiple persons spoke in the scene. We thereby proposed a novel speaker-aware speech enhancement (SASE) method that extracts speaker information using a “long short-term memory (LSTM)” method and a convolutional recurrent neural network (CRN) to match the specific speaker’s information. It was shown using a series of comprehensive experiments that only a few seconds of reference audio from the intended speaker suffice for the proposed SASE method to outperform the baseline systems.

Finally, the addition of a “self-attention” mechanism can further boost relevant speech-quality metrics. The methods were implemented in a multi-stage structure, referred to as

⁶ S. Pascual, A. Bonafonte and J. Serrà, "SEGAN: Speech Enhancement Generative Adversarial Network", *Proc. Interspeech 2017*, pp. 3642-3646, 2017.

multi-stage SA-TCN. The resulting solution performed well relative to other state-of-the-art systems in terms of speech enhancement and speech recognition scores.

3. Multi-style training

Multi-style training⁷ (MTR) is a technique we adopted to train our Kaldi speech recognition network to adapt to the various specific types of noise encountered in the battlefield. In a time delay neural network (TDNN) style neural network as Kaldi, MTR has been proven useful in better recognizing speech under specific types of pre-trained noise. Our results have shown that even with limited training data, MTR can significantly improve the speech recognition performance (Table I).

4. Language model enhancement

The language model (LM) in Kaldi is like a dictionary of words known to the speech recognition network. In the CARD project, we confirmed that the open-source Kaldi model's LM has lacked the medical terminologies involved in our use cases. We proceeded to enhance it by appending the model with all the relevant medical words using SRILM toolkit⁸ and confirmed improved speech recognition performance (Table I).

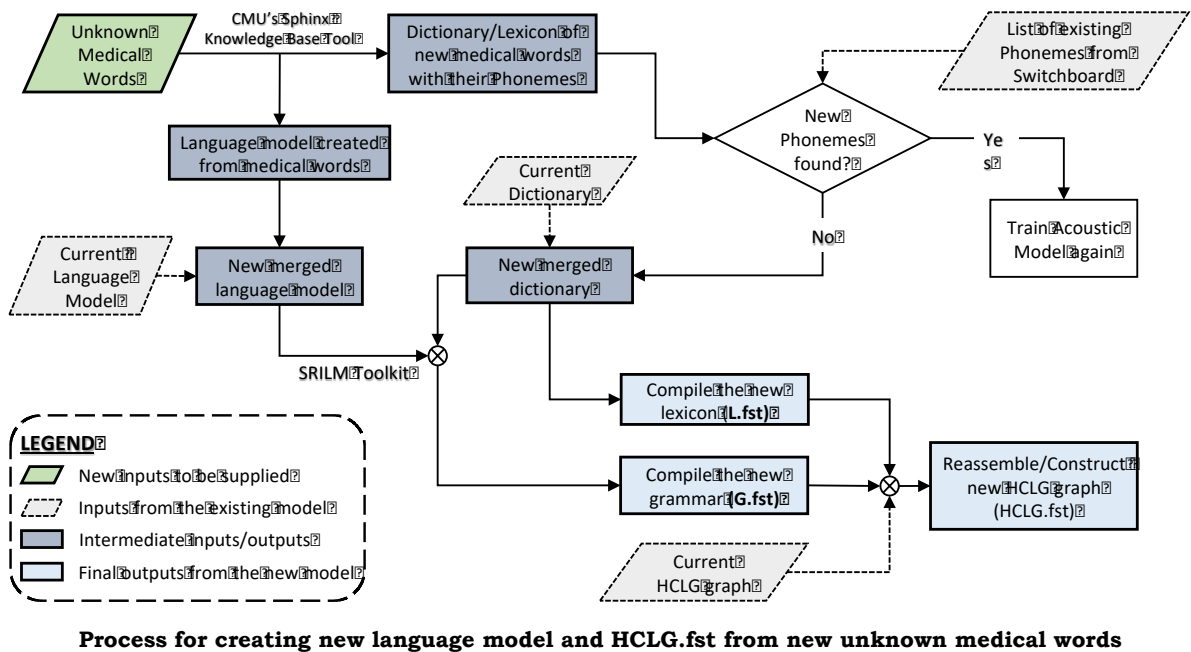


Figure 6. Language model enhancement process for Kaldi.

⁷ R. Lippmann, E. Martin and D. Paul, "Multi-style training for robust isolated-word speech recognition", *ICASSP'87. IEEE International Conference on Acoustics Speech and Signal Processing*, vol. 12, pp. 705-708, 1987.

⁸ A. Stolcke, "SRILM - an extensible language modeling toolkit", *7th International Conference on Spoken Language Processing ICSLP2002*, pp. 901-904, 2002.

5. Natural language processing for information extraction and post-hoc correction

After obtaining the ASR output with remaining errors in words and grammar, increasing academic attention has been placed on using natural language processing to 1) correct the errors and to 2) extract useful information from such unstructured texts.

For error correction, we explored “post-hoc correction” methods in recent literature that compare multiple possible sentence outputs from the ASR software and rank them based on observed probabilities from large training data sets made of correct sentences. Our preliminary findings suggest that, even though some word corrections have taken place, the unwanted side effects (incorrect corrections) can be substantial. Further research is needed in this area.

For information extraction, MetaMap⁹ is a key tool developed by the National Library of Medicine that has been widely used in biomedical information retrieval and data mining applications to obtain Unified Medical Language System (UMLS) Concept Unique Identifiers (CUIs) with corresponding textual descriptions. Our system used MetaMap with a predetermined group of medical concepts that led to information extraction for our specific domain.

6. Video analysis

In the end, the CARD project has opted to invest the majority of time in speech recognition. While we did investigate the use of YOLO – a convolutional neural network-based object detection method, we did not spend more time beyond confirming the basic software feasibility and the need to establish a military focused video data set for training YOLO.

7. Comparison with other related work

In conclusion, our combined efforts have achieved a medical word error rate (MWER) from 69.85% to 46.33% using the drill data, which is to-date the most accurate test data for assessing speech recognition performance in battlefield medical operations. In contrast, Facebook’s ASR system achieved about 20% WER¹⁰ on non-military noisy speech condition using about 13.7K hours of data. Google ASR, on the other hand, achieved about 40% WER with a medical data set¹¹.

Roadmap Development

Our roadmap development efforts identified three key areas most critical and promising for maturing the CARD methodology: errors introduced when multiple speakers are present in audio, the lack of additional non-audio inputs for correcting speech recognition errors, and the limited natural language processing algorithms adopted for error detection and correction. The finding guided us to focus our roadmap on enhanced speech separation methods for speech recognition,

⁹ A. R. Aronson and F.-M. Lang, "An overview of MetaMap: historical perspective and recent advances", *Journal of the American Medical Informatics Association*, vol. 17, no. 3, pp. 229-236, 2010.

¹⁰ Liu, Chunxi, et al. "Improving RNN transducer based ASR with auxiliary tasks." 2021 IEEE Spoken Language Technology Workshop (SLT). IEEE, 2021.

¹¹ Wang, Yongqiang, et al. "Transformer-based acoustic modeling for hybrid speech recognition." ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020.

joint audio-video recognition techniques, and a multitude of natural language processing techniques. Among them, our project to date has made most strides in the first and third aspects, i.e., speech enhancement and post-hoc correction.

The most critical next step for bringing the military ASR solution forward is the establishment of a significantly large-scale repository of audio and video recordings of military operations. It is critical that such data is adequately transcribed and labeled with their respective operating conditions. This will be the most valuable and necessary asset that will uniquely bootstrap a military ASR solution.

At the same time, the said audio and video repository need not occur in vacuum. By leveraging the US Forces' planned roadmap of smart phone deployment with the operational units, the audio and video recordings can be phased in with proper operation protocols. This not only will immediately begin to reduce the gap of documentation of battlefield medical operations, it will begin to accumulate the data needed to sustain a new industrial wave of military ASR and video analysis solutions.

What opportunities for training and professional development has the project provided?

The project has trained a significant number of graduate students across multiple areas, including nine PhD students (6 ECE, 3 Health Sciences), two MS students (1 ECE, 1 CS), and multiple undergraduate students involved in the drills. Amongst them, three PhD students have graduated and entered industry and academia, respectively. The team members also participated in several US and international conferences.

How were the results disseminated to communities of interest?

The results have been presented at multiple conferences and journals, including abstracts at MHSRS 2017, 2018 and full conference papers at INTERSPEECH 2019 & 2020, EUSIPCO 2021 and IJCNN 2021, and journal papers in JMIR mHealth and uHealth and IEEE Transactions on Audio, Speech and Language Processing.

What do you plan to do during the next reporting period to accomplish the goals?

This is the concluding period for the project.

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

The project has had major impacts on state-of-the-art understanding of speech recognition technology challenges for battlefield environment:

- Deep look into the question “is state-of-the-art speech solutions, either academic or commercial, ready for the battlefield?”
 - Produced systematic benchmarking of two most advanced open-source neural network-based speech recognition software by Mozilla (Deep Speech, end-to-end deep learning) and John Hopkins University (Kaldi, hybrid deep learning) with respect to battlefield noise of different types and intensity and concluded severe performance degradation with battlefield noise.

- Build foundations for DoD’s long-term success in speech recognition competencies:
 - Produced systematic assessment and analysis of speech recognition correctness’s dependencies on algorithms, processing, and data (training).
 - Derived approach to systematically assess specific speech recognition enhancement actions with measurable performance impacts
 - Developed modular enhancement methods to enable persistent performance enhancement for hybrid deep learning based speech recognition.
 - Integrated video analysis software for object and action analysis in recorded video.

- Build foundations for DoD’s long-term success in video analysis competencies:
 - Explored video analysis software for object and action analysis in recorded video.
 - Explored systematic process for DoD to continuously enhance video analysis with contents reflecting DoD operation needs and priorities.

- The research expects longer term impacts in:
 - Commercial, Civilian and Military
 - Commercial development in key technologies
 - Synergistic use in civilian EMS, homeland security and disaster recovery
 - New companies or consortium of companies on standard based integrated platform, from variety of data capture devices, mobile app, to cloud solutions
 - Integration with telemedicine, enhance the care and survivability of service members
 - Establish understanding of battlefield readiness of state-of-the-art technology
 - Establish foundations for long-term success for DoD speech recognition competencies

 - Academic
 - Speech recognition, natural language understanding, and video analysis in battlefield settings is a rich area for academic research
 - Large scale, distributed data collection (speech, video, sensors) and aggregation into individual patient data is a rich area of research

What was the impact on other disciplines?

Nothing to Report

What was the impact on technology transfer?

The technologies still have important challenges that required further research. The project has not explored technology transfer.

What was the impact on society beyond science and technology?

The project has advocated for the benefits and feasibility of a public-private partnership approach that can pave the way for industry success in military hands-free technologies. The approach emphasized the importance of embedding human in the solution, providing critical verification and error detection inputs that will help improve the technologies, while providing our soldiers with their documentation

needs. The concept will similarly impact the civilian medical rescue and first responder services, to help the society improve our ability to save lives, document such process for subject care and personnel training purposes.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

Nothing to Report

Actual or anticipated problems or delays and actions or plans to resolve them

Nothing to Report

Changes that had a significant impact on expenditures

Nothing to Report

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Nothing to Report

Significant changes in use or care of human subjects

Nothing to Report

Significant changes in use or care of vertebrate animals

Nothing to Report

Significant changes in use of biohazards and/or select agents

Nothing to Report

6. PRODUCTS:

- **Publications, conference papers, and presentations**

Journal publications.

- | |
|---|
| <ol style="list-style-type: none">1. "Speech enhancement using multi-stage self-attentive temporal convolutional networks", submitted to IEEE/ACM Transactions on Audio Speech and Language Processing.2. "Complete and resilient documentation (CARD) for operational medical environments leveraging hands-free technology in a systems approach," submitted to Journal of the American Medical Informatics Association. |
|---|

Books or other non-periodical, one-time publications.

Nothing to Report.

Other publications, conference papers and presentations.

1. Ju Lin, Sufeng Niu, Zice Wei, Xiang Lan, Adriaan J Wijngaarden, Melissa C Smith, Kuang-Ching Wang, “Speech enhancement using forked generative adversarial networks with spectral subtraction,” in Proceedings of Interspeech 2019, 2019.
2. Ju Lin, Sufeng Niu, Adriaan J van Wijngaarden, Jerome L McClendon, Melissa C Smith, Kuang-Ching Wang, “Improved Speech Enhancement using a Time-Domain GAN with Mask Learning,” in Proceedings of Interspeech, 2020.
3. J Lin, AJ van Wijngaarden, KC Wang, MC Smith, “Speaker-aware Speech Enhancement with Self-attention,” in Proceedings of EUSIPCO (European Signal Processing Conference), 2021.
4. Snigdhaswin Kar, Prabodh Mishra, Ju Lin, Min-Jae Woo, Nicholas Deas, Caleb Linduff, Sufeng Niu, Yuzhe Yang, Jerome McClendon, D. Hudson Smith, Melissa Smith, Ronald Gimbel and Kuang-Ching Wang, “Systematic Evaluation and Enhancement of Speech Recognition in Operational Medical Environments,” in Proceedings of IJCNN (International Joint Conference on Neural Networks), 2021.
5. South Carolina ETV article: New research to help injured troops using technology https://www.scetv.org/stories/health/2020/new-research-help-injured-troops-using-technology?fbclid=IwAR3lN0hYeqMNqZ_8rEGocqA07exvnAnnjc0eJDhcdmPGRnmESeEh2lPsgSk

- **Website(s) or other Internet site(s)**

Nothing to Report.

- **Technologies or techniques**

Nothing to Report.

- **Inventions, patent applications, and/or licenses**

Nothing to Report.

- **Other Products**

- Speech data sets
 - Switch Board audio data
 - Common Voice audio data
 - Libris Speech
 - Medical drill audio and video recordings
- Noise augmentation
 - 14 types of battlefield noises
 - Intermittent and continuous noises

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS:

What individuals have worked on the project?

Name:

Kuang-Ching Wang

Project Role:

Principal Investigator

Research Identifier:

0000-0002-5675-7104

Contribution to Project:

Dr. Wang is responsible for all aspects of the research study.

Name: Ronald Gimbel
Project Role: Co-PI
Research Identifier: 0000-0001-8185-4013
Contribution to Project: Dr. Gimbel supports Dr. Wang and Dr. Griffin in qualitative research, co-coordinating drill #1, and communication with the Palmetto Health Emergency Medicine team.

Name: Sarah Griffin
Project Role: Investigator
Research Identifier: 0000-0003-4820-3985
Contribution to Project: Dr. Griffin will lead focus groups and interviews to glean insight on operational medical workflow.

Name: Melissa Smith
Project Role: Investigator
Research Identifier: 0000-0003-0798-8536
Contribution to Project: Dr. Smith and her graduate students will lead technology enablement activities.

Name: Ashwin Srinath
Project Role: Investigator
Research Identifier: 0000-0003-4341-9840
Contribution to Project: Dr. Srinath will oversee integration of the CPU/GPU platforms.

Name: Sufeng Niu
Project Role: Graduate Student
Contribution to Project: Mr. Niu will support Dr. Smith's work on speech recognition.

Name: Yuzhe Yang
Project Role: Graduate Student
Contribution to Project: Ms. Yang will focus on experiment design and publication.

Name: Xiang Lan
Project Role: Graduate Student
Contribution to Project: Mr. Lan will support Dr. Smith's work on video and speech recognition.

Name: William Gerard, MD
Project Role: Co-Investigator
Contribution to Project: Dr. Gerard oversaw all aspects of the study at Prisma Health-Midlands.

Name: Stephen Shelton, MD
Project Role: Investigator
Contribution to Project: Dr. Shelton worked alongside Dr. Gerard and the Emergency Department physicians at Prisma Health-Midlands.

Name: Christopher Gainey, MD
Project Role: Investigator
Contribution to Project: Dr. Gainey worked alongside Dr. Gerard and the Emergency Department physicians at Prisma Health-Midlands.

Name: Kelly Hawsey, RN
Project Role: Investigator
Contribution to Project: Mrs. Hawsey assisted with logistics of the study at the Prisma Health-Midlands.

Name: Jerome McClendon
Project Role: Post-Doc
Contribution to Project: Dr. McClendon assisted with multiple aspects of the CARD technologies research.

Name: Minjae Woo
Project Role: Graduate Student
Contribution to Project: Mr. Woo supports Drs. Gimbel and Griffin in activities related to the proposed qualitative research, statistical analysis and outcome analysis, and issues related to the drills and workflow. The student will also develop the basic model that controls simulation of battlefield noise.

Name: Caleb Linduff
Project Role: Graduate Student
Contribution to Project: Mr. Linduff will focus on platform integration and support Dr. Wang in the area of how to integrate the data into the Cerner electronic health record.

Name: Qing Wang
Project Role: Graduate Student
Contribution to Project: Mr. Wang will support Dr. Smith's work on video and speech recognition.

Name: Ju Lin
Project Role: Graduate Student
Contribution to Project: Mr. Lin will support Dr. Smith's work on video and speech recognition.

Name: Snigdhaswin Kar
Project Role: Graduate Student
Contribution to Project: Mr. Kar will support Dr. Wang's work on benchmark experiments and publication.

Name: Prabodh Kumar Mishra
Project Role: Graduate Student
Contribution to Project: Mr. Mishra will support Dr. Wang's work on benchmark experiments and publication

Name: J N Deas
Project Role: Undergraduate Student
Contribution to Project: Mr. Deas will support Dr. Wang's work on benchmark experiments and publication

Name: Brian Helsel
Project Role: Graduate Student
Contribution to Project: Mr. Helsel engaged his experience in working with big data from the electronic health record systems.

Name: Anamal Haque
Project Role: Graduate Student
Contribution to Project: Mr. Haque investigated machine learning methods used in medical applications.

Name: Jacob Lipton
Project Role: Undergraduate Student
Contribution to Project: Mr. Lipton focused on the investigation of open source machine learning software for text analysis.

Name: Karen Edwards
Project Role: Research Associate
Research Identifier: 0000-0003-3584-0328
Contribution to Project: Mrs. Edwards assisted with regulatory/IRB approvals as well as the approval from the HRPO military secondary review board. She will assist with various reports and logistical items.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to Report.

What other organizations were involved as partners?

Organization Name: Prisma Midlands

Location of Organization: Columbia, South Carolina

Partner's contribution to the project: Prisma Midlands provided staff and facilities for this project. Staff included Emergency Medicine physicians and nurses, EMTs, Flight nurses and administrative staff. Their Simulation Center was used for the drills that were completed. Physicians included Christopher Gainey, Steve Shelton and William Gerard. Prisma administrator was Kelly Hawsey.

8. SPECIAL REPORTING REQUIREMENTS:

COLLABORATIVE AWARDS: N/A

QUAD CHARTS: See Appendix

9. APPENDICES:



Complete and Resilient Documentation (CARD) for Operational Medical Environments
 W81XWH-17-C-0238
 Award Amount: \$1,644,510

Log Number: DM160374
 PI: Kuang-Ching Wang
 ORG: Clemson University

Study/Product Outcome/Aim(s)

- Our objectives are:
 - o Identify the medical workflow and information needs of battlefield medics and air transport nurses caring for wounded warriors in an operational environment.
 - o Develop the platform for hands free EHR data documentation robust to the environment.
 - o Develop natural language processing solution for multi-stage information extraction.
 - o Develop systematic assessment instruments based on medical workflow efficacy.

Approach

The research targets long term feasibility of hands-free documentation of injury in operational environments from two fronts: Medical and Technological.

- On the medical front:
 - o Focus groups/interviews
 - o Simulation drills (battlefield, en route, field hospital care)
- On the technology front:
 - o Platform for flexible integration (recording devices, compute devices, and EHR)
 - o Decoupled recording and processing + persistent performance improvement



Accomplishments: 1. focus groups with EMS personnel, 2. capture & speech recognition system integrated, 3. systematic noise impact study (for speech recognition), 4. two medical drills, 5. medical info extraction from speech

Goals/Milestones

- CY17-18 Goals – Medical Workflow and Platform Development
 - Focus Groups, Drill #1
 - Platform Integration
 - Speech Recognition
 - CY18-20 Goals – Qualitative Assessment and Roadmap Development
 - Information Extraction and integration with EHR
 - Video Digest
 - Drill #1a-b, #2
 - Workflow Assessment
- Comments/Challenges/Issues/Concerns
- Base system and study to date identifies challenges & promising opportunities for further development of robust solutions.
- Budget Expenditure to Date
 Projected Expenditure: \$1,644,510
 Actual Expenditures: \$1,635,203 @30JUNE2021

Timeline and Cost

Activities	17	18	19	20	21
Medical Workflow Studies					
Platform development					
Assessment					
Roadmap development					
Estimated Budget (\$K)				\$893k	\$751k

Simulation Drill #2 Photographs

Transport scenario with amputee patient



Hospital scenario with a gunshot wound patient



Battlefield scenario with a concussion



Highlights from Simulation Drill #2 on December 2, 2019

Simulation drill #2 was the last drill before this project concludes. It consisted of the following components and emphases:

- Full CARD solution on the NVIDIA TX2-based embedded platform
- Full Kaldi/TensorFlow based speech recognition software with speech enhancement, medical language model, preliminary NLP post-processing, and TCCC card entry
- Monitoring dashboard with live streamed status and speech, video, and TCCC outputs of all (6) active TX2 platforms
- Extended medic scripts covering more data fields in the TCCC card
- Tests under multiple conditions: 1) high vs. low vs. no noise, 3) single vs. multiple speaker, 4) regular vs. speed-up pace of speech
- Improved video camera attachment to medics and recording of point-of-injury surroundings

LEARNING OBJECTIVES

- Learn and discuss the pressing needs and challenges for health record collection in operational environments.
- Understand the state of deep learning based speech recognition and its potential use in health record collection in battlefield.
- Understand the range of GPU compute technology and its use in multiple form factors.
- Understand the major study thrusts to develop, assess and deploy the CARD system.

ABSTRACT

- The project's mission is to demonstrate and validate hands free electronic health data entry that:
 - has reliability in a noisy environment
 - alleviates disruption of care for documentation
 - prevents loss of documentation
- We have proposed a system-oriented approach that:
 - enables resilient hands-free data collection
 - preserves complete documentation
 - presents timely information useful to the medical operation
- We will show that existing technology is adequate if the data record is decoupled from documentation entry
- CARD will show uninterrupted medical workflow by:
 - Preserving complete documentation
 - Presenting timely data relevant to the roles of care

BACKGROUND

- Existing Natural Language Processing (NLP) solutions are not adequate in battlefield
 - Results in loss of critical information
- Concerns with existing data capture solutions
 - Disrupts medics from delivering care
 - Requires medic-worn devices to be both light-weight and powerful

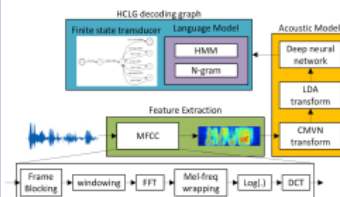
METHODS

- The project team includes technology researchers, military health system experts, and emergency medical physician leaders
- The project is integrated with the Cerner electronic health record (EHR) system
 - Data will be validated at initial and final stages
- An accredited medical simulation center will be used
 - Multiple levels of care (Role 1, 2, 3) scenarios with military-oriented EMS staff and physicians



- Three paths will be taken in parallel
 - Work flow identification and information requirements
 - Interviews and focus groups will be used
 - Emergency medical physicians, EMS personnel, and medical transport nurses
 - Hardware research
 - GoPro and non-volatile storage
 - GPU platforms
 - Cerner EHR
 - Software Research
 - Building a program and library for NLP based speech recognition and video classification software
- The open architecture assures CARD to be a flexibly extensible and long-term evolvable platform

METHODS (CONT')



RESULTS AND CONCLUSION

The project has been recommended for funding in February 2017. The project's initial focus is on clinical interviews, simulation center drills, platform integration, and speech recognition algorithm development.

Alpha benchmarks for the system are expected by Spring 2018.

AFFILIATIONS

- ¹Office of the Chief Medical Information, OTSG, U.S. Army
- ²Department of Public Health Sciences, Clemson University
- ³Department of Electrical Engineering, Clemson University
- ⁴Department of Computing, Clemson University
- ⁵Palmetto Health System, Columbia, SC

DISCLAIMER

This work is recommended for funding by the U.S. Department of Defense Defense Medical Research and Development Program under Grant No.WB1XWH-16-R-MS14. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Defense.

Complete and Resilient Documentation (CARD) for Operational Medical Environments

March 27, 2018

- ❖ Organization / key partners / institutes: **Clemson University**
- ❖ Award #: W81XWH-17-C-0238
- ❖ Award Mechanism: Contract
- ❖ Solicitation: BAA, Health Information Technologies and Informatics Hands-Free Electronic Health Record Data Entry Initiative (HFEHRI)
- ❖ Principal Investigator: Kuang-Ching Wang, PhD
 - ❖ Key Sub-Awards: Bill Gerard, MD (Palmetto Health System)
- ❖ Total Cost/Budget: \$1,644,510
- ❖ Period of Performance: 9/25/17-12/24/19
- ❖ Grants / Contract Officer Representative: Dr. Prem Yadav
- ❖ Grants / Contract Specialist: Jeanette Wolfe
- ❖ Related government funding: n/a

PI: Kuang-Ching Wang

Org: Clemson University

Award Amount: \$1,644,510

Study/Product Aims

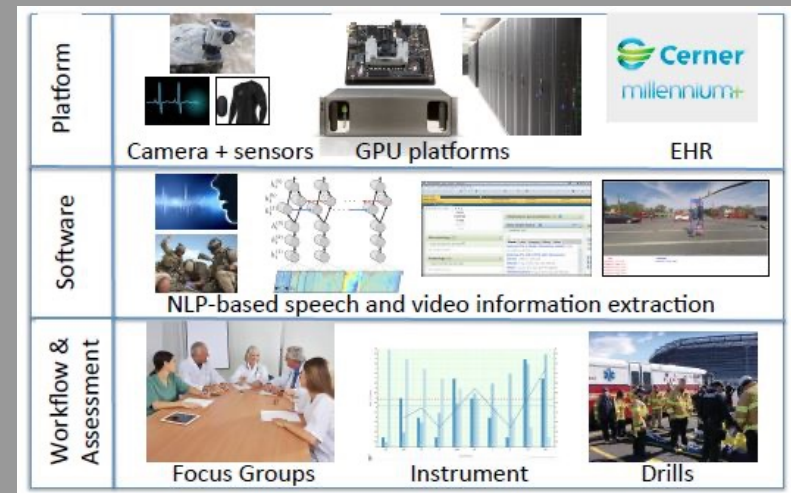
Our objectives are:

- **Identify the medical workflow** and information needs of battlefield medics and air transport nurses caring for wounded warriors in an operational environment.
- **Develop the platform** for hands free EHR data documentation robust to the environment.
- **Develop natural language processing solution** for multi-stage information extraction.
- **Develop systematic assessment instruments** based on medical workflow efficacy.

Approach

The research targets long-term feasibility of hands free documentation of injury in operational environments from two fronts: **Medical** and **Technological**.

- **On the medical front:**
 - Focus groups/interviews
 - Simulation drills (battlefield, en route, field hospital care)
- **On the technology front:**
 - Platform for flexible integration (recording devices, compute devices, and EHR)
 - Decoupled recording and processing + persistent performance improvement.



The study is structured into three parallel thrusts: platform development, software development, and workflow & assessment.

Timeline and Cost

Activities	CY	17	18	19
Medical workflow studies			■	■
Platform development		■	■	
Assessment			■	■
Roadmap development				■
Estimated Budget (\$K)			\$893k	\$751k

Goals/Milestones

CY17-18 Goals – Medical Workflow, Platform Development

- Focus Groups, Drill #1
- Platform Integration
- Speech Recognition

CY18-19 Goals – Qualitative Assessment, Roadmap Development

- Video Digest
- Cerner Integration
- Drill #2
- Workflow and EHR Assessment

Comments/Challenges/Issues/Concerns

- There are no challenges, issues or concerns at this time.

Budget Expenditure to Date

Projected Expenditure: \$1,644,510

Actual Expenditures: \$95,157.23 @28FEB2018

- ❖ Purpose: This research project addresses one common perception of hands free electronic health record data collection for the operational environment
 - ❖ Hands free data entry requires speech recognition, but ...
 - ❖ It is too hard to do in the field with all the noise, movement, and harsh elements
- ❖ Methodology
 - ❖ Decoupled recording and recognition
 - ❖ Robust recording in the field + Just in time speech recognition + perpetual archive
 - ❖ Medical leads the way
 - ❖ Focus group studies with EMS physicians, helicopter transport nurses with prior military medical experiences + EMS drills + systematic quality assessment
 - ❖ Platform for integration
 - ❖ All components are evolvable: Data input devices, compute HW/SW, storage, EHR interface
 - ❖ Resilience and usability rooted in platform design and data lifecycle
 - ❖ Focus on system architecture & data interfaces, instead of individual component performance
 - ❖ Creation of military-oriented training data set for speech and video recognition
- ❖ Expected Outcomes
 - ❖ Medical workflow and requirements
 - ❖ CARD platform prototype hardware and software
 - ❖ Assessment instruments and results




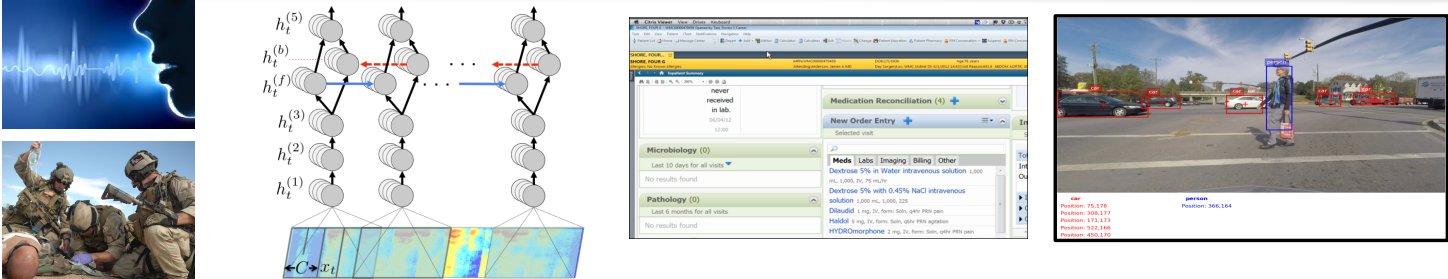

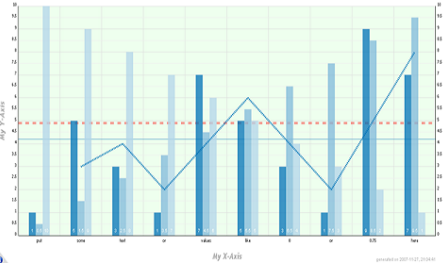

- ❖ Hypothesis(es) / Research Question
 - ❖ The key objective for the solution is two-fold
 - ❖ Complete and resilient documentation (CARD)
 - ❖ Timely data that enhances medical workflow
 - ❖ The key concerns for the solution differ based on roles of care
 - ❖ Point of injury (medics) – record, confirm correct capture
 - ❖ Transport/en route care (nurses) – record further, confirm correct capture, communicate advance stats downstream
 - ❖ Field Hospital (physicians) – review assessment and treatment history since point of injury
 - ❖ Recorded data can be in multiple forms linked to a patient's EHR record
 - ❖ Original audio/video recording
 - ❖ Structured text data via speech recognition and natural language understanding
 - ❖ The former assures complete and resilient documentation. The latter assures useful enhancement to medical care
 - ❖ Speech recognition performance can be persistently improved given
 - ❖ A sizeable archive of recordings for training deep-learning based software
 - ❖ Decoupled “human-audit” and “re-train” process using the archive.

- ❖ **Specific Aims / High Level Objectives**
 - ❖ **Aim 1: Medical Workflow**
 - ❖ Identify requirements and operational conditions from physicians and nurses
 - ❖ Study medical rescue standards and deviations in operational environments
 - ❖ **Aim 2: Platform**
 - ❖ Integrate, end-to-end, cameras, GPU compute devices, storage, to Cerner EHR
 - ❖ Develop and integrate speech recognition, video digest solutions
 - ❖ Explore a multitude of data interface and management options
 - ❖ **Aim 3: Assessment**
 - ❖ Experienced physician/nurse/EMS drills
 - ❖ Surveys for workflow efficiency
 - ❖ **Aim 4: Technology roadmap development**

- ❖ **Impact:** The proposed work has strong potential to advance the standard of care for both injured service members and the general public. The impacts include evolved standards for reliable documentation of injury and, subsequently, better care delivery and benefits management. They form critical foundations for enhanced field care, telemedicine, soldier survivability, and benefits management.

Research Design / Methods

The CARD Project

<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Platform</p>	   <p>Camera + sensors GPU platforms EHR</p>
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Software</p>	 <p>NLP-based speech and video information extraction</p>
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Workflow & Assessment</p>	   <p>Focus Groups Instrument Drills</p>

Focus Group Design

- **Population:** ED Physicians, Emergency Medical Service (EMS) personnel and medical transport nurses
- **Quantity:** Two focus groups with 4-6 participants for each of the occupational groups
- **Design:** Based upon the medical information that should be collected throughout the care, and the appropriate time and form to record them into the HER
- **Results:** Recorded, transcribed and analyzed using the traditional grounded theory approach. An analysis using a constant comparative method to identify initial codes (open coding) then examine the interrelationships within and between codes and code groups (axial coding) will be conducted.
- A formal report with recommendations will be generated to advise the CARD technology development team and will be delivered to USAMRMC; a scientific manuscript on the outcome will also be crafted

Drill #1 Design

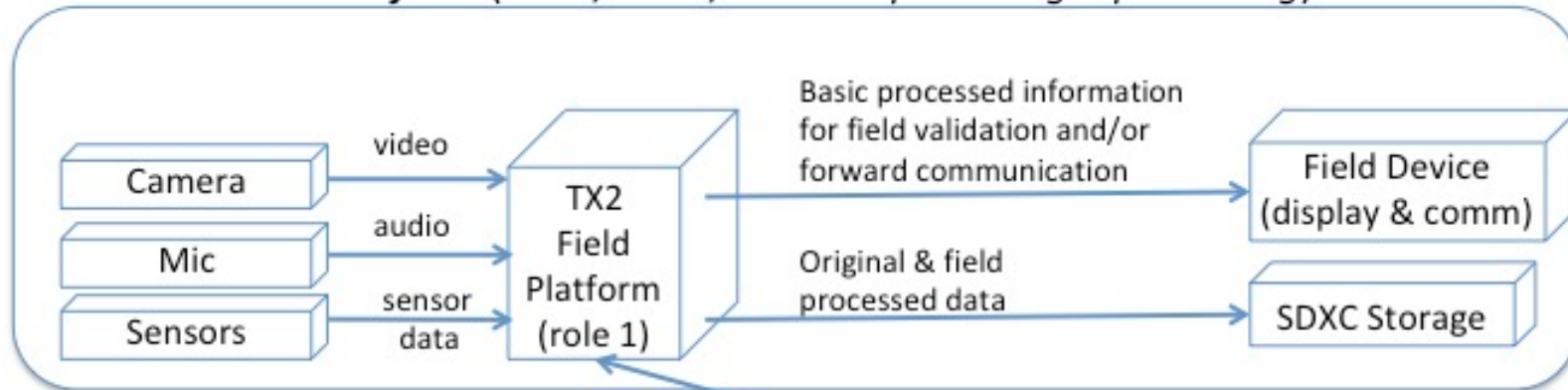
- ❖ **Simulated injuries**
 - Below knee amputation with arterial bleeding secondary to an improvised explosive device injury.
 - Gunshot wound through and through chest with respiratory distress.
 - Blast injury with concussion (altered mental status) and tympanic membrane perforations.
- ❖ **Simulated settings**
 - Battlefield, with gunshot noise
 - Air transport, with helicopter noise
 - Military field hospital

Research Design / Methods

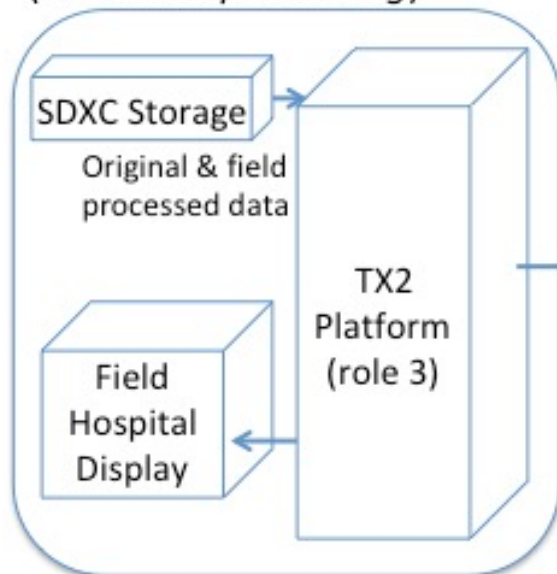
The CARD Platform



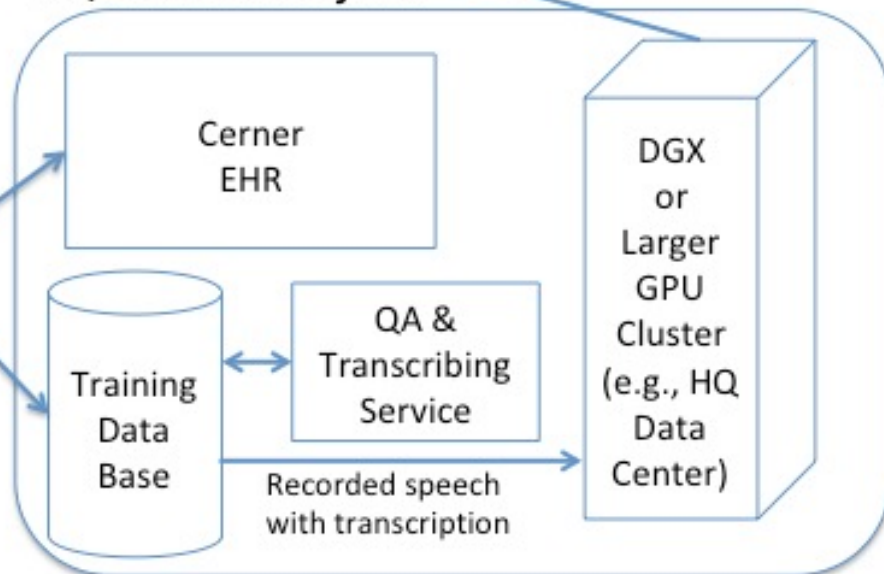
Role 1 Field Platform (video/audio/sensor capture + light processing)

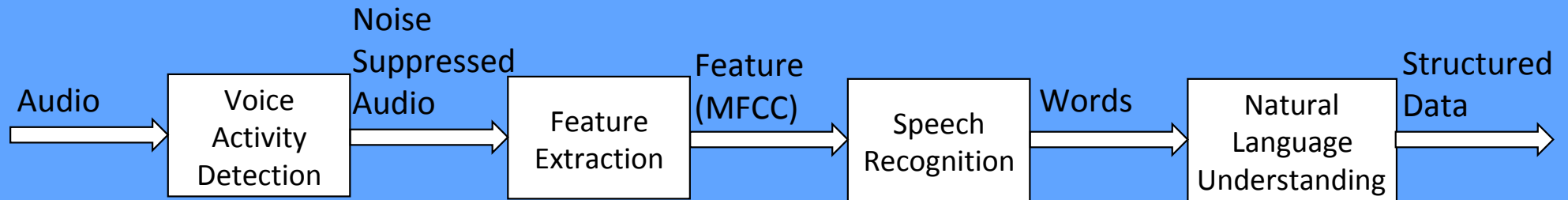


Role 3 Field Hospital Platform (enhanced processing)



HQ Backend Platform





➤ Current focus – integration of open source components:

- ❑ Noise suppression – Google WebRTC Voice Activity Detection

<https://webrtc.org/>

- ❑ Deep learning neural networks – Mozilla/Baidu Project DeepSpeech

<https://github.com/mozilla/DeepSpeech>

Above delivers state of the art performance in civilian settings, but significantly degrades with battlefield noise. Focus now zooms in on new noise-tolerant deep learning solutions.

➤ Performance driven tuning

- ❑ Speech recognition performance (Word Error Rate, WER) vs.
- ❑ Speech recognition (inference) time vs.
- ❑ Neural network training time vs.
- ❑ Data memory requirement

❖ Simulation Center Drill #1

- ❖ 1.2 hours video/audio recording in three settings, fully transcribed
- ❖ Initial speech recognition performance ~20% Word Error Rate (comparable to DeepSpeech reported performance)



❖ Focus Group

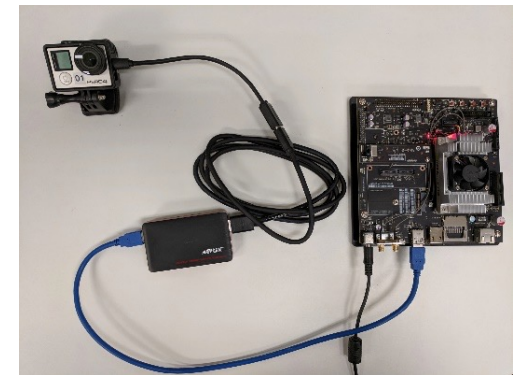
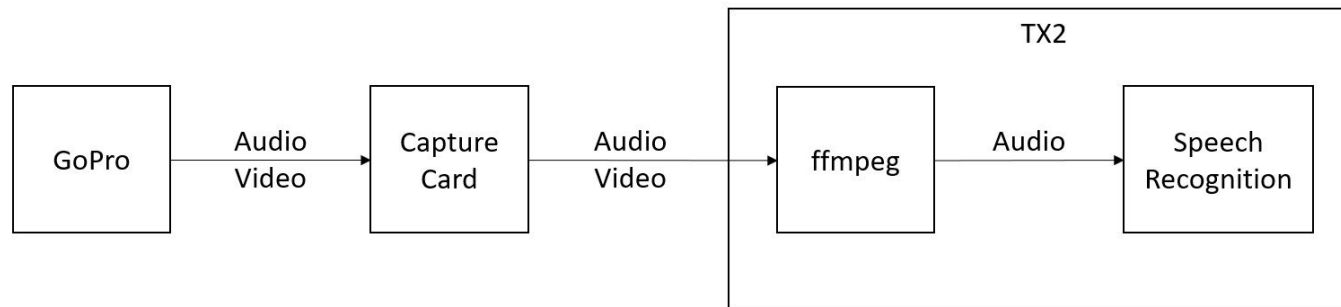
❖ Recruitment demographics

	TOTAL n	Experience <10 years	Experience >10 years
EMS	6	17%	83%
Transport RNs	11	0%	100%
ED MDs	9	33%	67%
TOTAL	26		

❖ Significant inputs to technology team

- ❖ Verbal and non-verbal communication; sometimes many people talking, sometimes multiple patients
 - ❖ Strong desire for sending advance stats to physicians
 - ❖ Strong opinions on camera mounting positions, on/off control
- ❖ A formal report with recommendations will be generated to advise the CARD technology development team and will be delivered to USAMRMC
- ❖ A scientific manuscript on the focus group design, recruitment, interviews, transcription, analysis and operation recommendations will also be crafted.

- ❖ Speech recognition
 - ❖ Complete Alpha development, working on algorithm variations for better performance, benchmarking and investigating speed-up options
- ❖ Platform integration
 - ❖ Completed – Camera + video capture + TX2 GPU board + software



- ❖ Cerner integration
 - ❖ Account setup for Cerner FHIR sandbox

❖ COST

- ❖ What is the Money(\$) spent to date? \$95,157.23
- ❖ Is burn rate on schedule? Yes
- ❖ Have you spent too much or too little; and why?
This project is financially on track.

❖ SCHEDULE

- ❖ Is the project on schedule, behind schedule, what delays?
This project is on schedule.

❖ PERFORMANCE

- ❖ What tasks have you successfully completed on your SOW?
 - ❖ Task 1.1: Focus Group/Interviews
 - ❖ Task 1.2: First EMS Drill
- ❖ What tasks are at risk ? There are no tasks at risk.

GANTT Chart



	Q1	Q2	Q3
Specific Aim I- Medical Workflow Studies			
Task 1.1 Focus Groups/Interviews			
Focus groups/interviews design and recruitment (100%)			
Focus groups/interviews meet (100%)			
Transcription, analysis, operation recommendation (50%)			
IRB Approvals (Clemson, PHS, HRPO) (100%)			
Task 1.2: First EMS Drill			
Drill #1 design, preparation and execution (100%)			
Specific Aim 2- Platform Development			
Task 2.1: Platform Integration			
Acquisition of components (10%)			
Platform and data interface integration (100%)			
Speech and video software integration (50%)			
Task 2.2: Speech Recognition & Video Digest			
Basic speech recognition development (100%)			
Training data preparation & training (90%)			

Specific Aim 1 – Medical Workflow Studies

Task 1.1: Focus Group/Interviews

Transcription, analysis and operation recommendation of focus groups – to be completed by 31MAR2018

Specific Aim 2 – Platform Development

Task 2.1 – Platform Integration

Acquisition of components – Technical supplies were purchased for Drill #1, further acquisition of GPU system will be done by 31MAR2018 for software testing.

Platform and data interface integration – Based on the GPU system, we will verify all data interfaces for end to end audio and video data storage and processing - to be completed by 31MAR2018

Cerner interface validation – We will develop the CARD-Cerner interface using the FHIR Cerner sandbox - to be completed by 31MAR2018

Task 2.2 – Speech Recognition & Video Digest

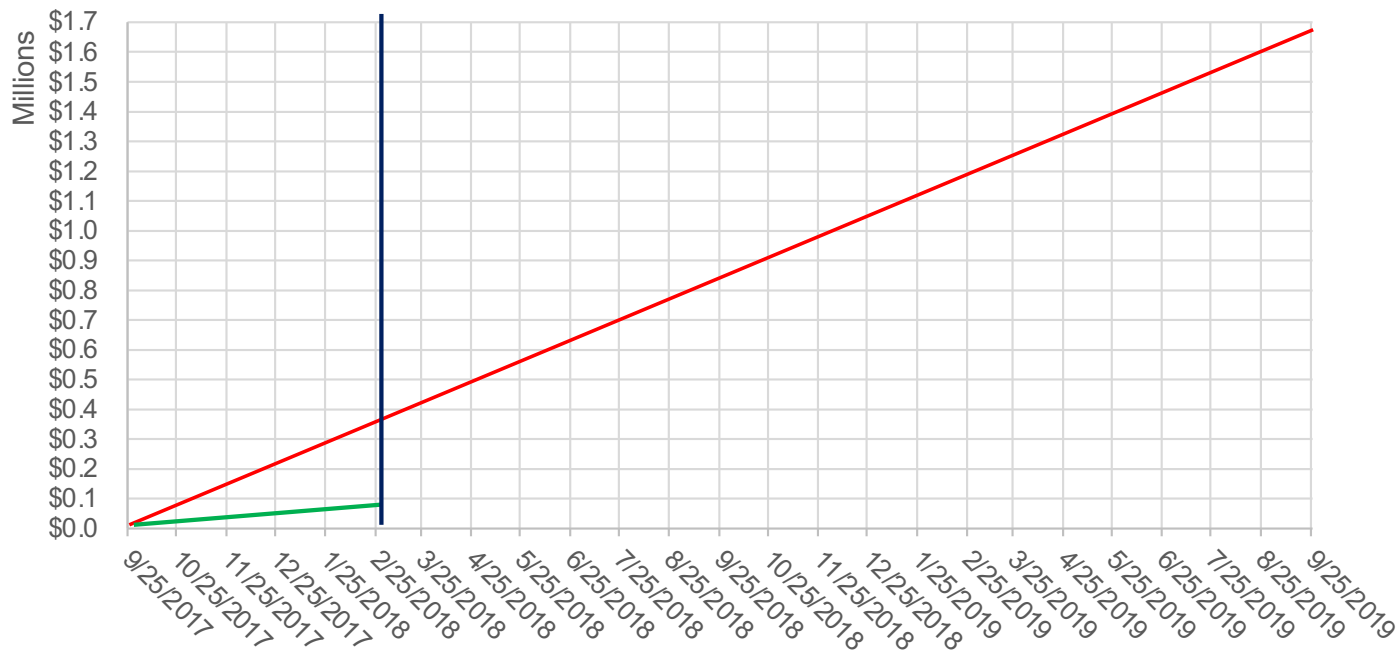
Speech Recognition development (months 1-6) – This task refers to the end-to-end speech recognition integration with the platform which will be completed by 31MAR2018. We will investigate further noise reduction techniques to make speech recognition more robust in a highly noisy environment till 30JUNE2018.

Data preparation & training (months 4-9) - We will train the speech recognition model based on Deep Learning techniques. We will try different parameter settings and different model structures. We will generate additional speech data for training – to be completed 30JUNE2018.

- ❖ Budget Expenditure to-date:
 - ❖ Projected Expenditure: \$1,644,510
 - ❖ Actual Expenditure: \$95,157.23
 - ❖ Anticipated adjustments on remaining periods – No known adjustments at this time

- ❖ Speak to expected vs. actual burn rates
 - ❖ Are you spending money too fast or too slow, and why? This study is financially on track.

Burn Rate



Contracting Overview



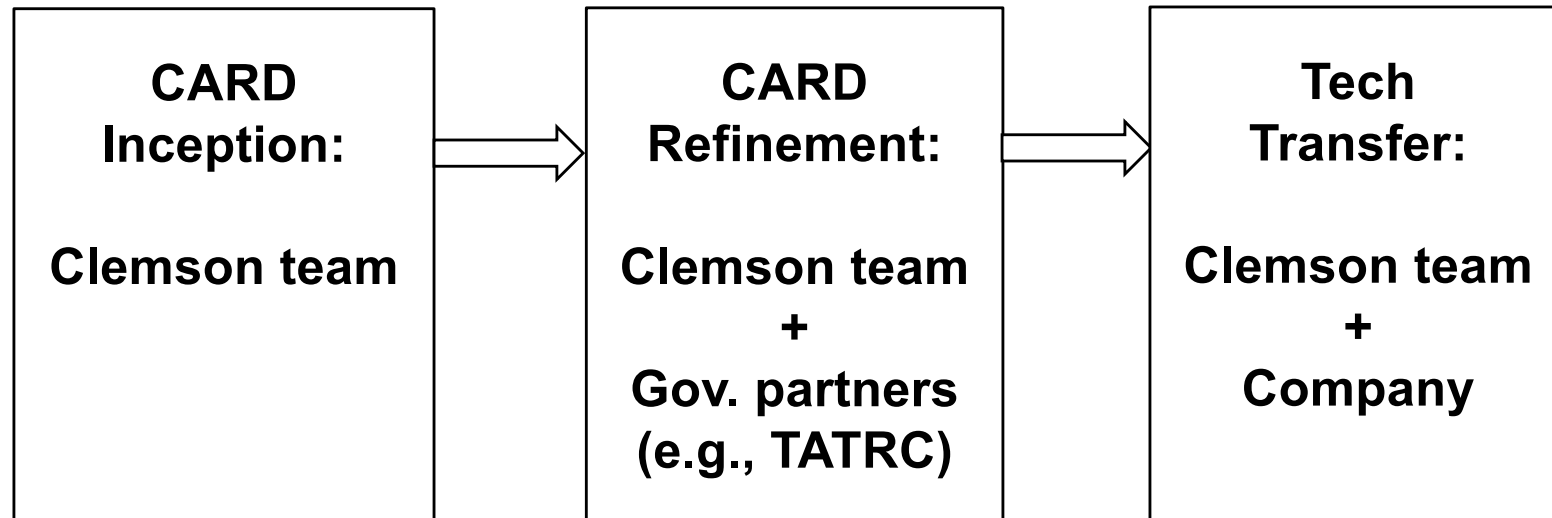
- ❖ HIGHLIGHT tasks that have deviated from original proposal or SOW.
 - ❖ There are no tasks that have deviated from the original proposal or SOW.

Deliverables & Data Requested	Details/Format	Data Rights	Delivery Schedule
Simulation Drill #1 audio/video data	Recorded data and transcription, part of a data set to be used and published for military operational speech recognition training.	n/a	12/11/17
Quarter #1 Technical Progress Report and Quad Chart	A brief description of overall progress to date for the reporting period 9/25/17-12/24/17	n/a	12/24/17
Quarter #2 Technical Progress Report and Quad Chart	A brief description of overall progress to date for the reporting period 12/25/17-3/24/18	n/a	3/24/18
Focus Groups	<p>A manuscript on focus group design, recruitment, interviews, analysis and operation recommendations will be crafted, including the analysis of recordings as provided by transcripts received.</p> <p>A formal report with recommendations will be generated to advise the CARD technology development team and will be delivered to USAMRMC.</p>	n/a	TBD

Anticipated Impact as an Outcome of Research



- ❖ The research expects significant impacts in:
 - ❖ Commercial, Civilian and Military
 - Commercial development in key technologies
 - Synergistic use in civilian EMS, homeland security and disaster recovery
 - New companies or consortium of companies on standard based integrated platform, from variety of data capture devices, mobile app, to cloud solutions
 - Integration with telemedicine, enhance the care and survivability of service members.
 - ❖ Academic
 - ❖ Speech recognition, natural language understanding, and video uses in battlefield settings is a rich area for academic research
 - ❖ Large scale, distributed data collection (speech, video, sensors) and aggregation into individual patient data is a rich area of research



- Military & academic research in speech, NLU & clinical info extraction & use in operational environments
- Civilian applications for emergency medicine & homeland security

- ❖ As it completes the 6th month, the project is on track on all proposed tasks.
- ❖ The anticipated outcomes and their status/outlook are as follows:
 - ❖ Focus group study – study completed, publication in progress
 - ❖ Platform development and concept demonstration – basic speech recognition completed, performance study in progress, natural language understanding and video digest to begin
 - ❖ (2nd) simulation drill & quality assessment
- ❖ The project's aim is not a point product. Instead, its aims are on:
 - ❖ Confirmation of the CARD concept's clinical opportunities and challenges
 - ❖ Development of the platform's architecture and interfaces
 - ❖ Demonstrating the extensible nature of the platform components
 - ❖ Assessing the CARD design's impact to the medical workflow

BACK-UP SLIDES

- ❖ Palmetto Health System IRB Approval – 10/25/17
- ❖ Clemson University IRB Approval – 10/26/17
- ❖ HRPO Military Secondary Review Approval – 11/17/17

❖ Poster presentation

- ❖ Wang, K., Gimbel, R., Smith, M., Ziolkowski, M., Jennings, W., Shelton, S., Gerard, W., Hawsey, K. Complete and Resilient Documentation (CARD) for Operational Medical Environments. 2017 Military Health System Research Symposium (MHSRS), Kissimmee, FL, August 27-30, 2017

❖ Intended future Publications / Presentations

- ❖ Griffin, S., Gimbel, R., Wang, K., Edwards, K., Gerard, W., Shelton, S., Hawsey, K. Complete and Resilient Documentation (CARD) for Operational Medical Environments: Knowledge Gained from Focus Groups Regarding the Use of GoPro® Equipment in Military and Civilian Medical Environments, 2018 MHSRS, Location and date TBD
- ❖ A formal report with recommendations will be generated to advise the CARD technology development team and will be delivered to USAMRMC
- ❖ A scientific manuscript regarding the focus group design, recruitment, interviews, transcription, analysis and operation recommendations will also be crafted.

Complete and Resilient Documentation (CARD) for Operational Medical Environments

January 22, 2019

- ❖ Organization / key partners / institutes: Clemson University
- ❖ Award #: W81XWH-17-C-0238
- ❖ Award Mechanism: Contract
- ❖ Solicitation: BAA, Health Information Technologies and Informatics Hands-Free Electronic Health Record Data Entry Initiative (HFEHRI)
- ❖ Principal Investigator: Kuang-Ching Wang, PhD
 - ❖ Key Sub-Awards: Bill Gerard, MD (Palmetto Health System)
- ❖ Total Cost/Budget: \$1,644,510
- ❖ Period of Performance: 9/25/17- 9/24/19
- ❖ Grants / Contract Officer Representative: Dr. Prem Yadav
- ❖ Grants / Contract Specialist: Jeanette Wolfe
- ❖ Related government funding: n/a

Updated: (12/31/18)

Study/Product Outcome/Aim(s)

Our objectives are:

- **Identify the medical workflow** and information needs of battlefield medics and air transport nurses caring for wounded warriors in an operational environment.
- **Develop the platform** for hands free EHR data documentation robust to the environment.
- **Develop natural language processing solution** for multi-stage information extraction.
- **Develop systematic assessment instruments** based on medical workflow efficacy.

Approach

The research targets long-term feasibility of hands free documentation of injury in operational environments from two fronts: **Medical** and **Technological**.

On the medical front:

- Focus groups/interviews
- Simulation drills (battlefield, en route, field hospital care)

On the technology front:

- Platform for flexible integration (recording devices, compute devices, and EHR)
- Decoupled recording and processing + persistent performance improvement

Platform	<p>Camera + sensors GPU platforms EHR</p>
Software	<p>NLP-based speech and video information extraction</p>
Workflow & Assessment	<p>Focus Groups Instrument Drills</p>

Accomplishments: 1. focus groups with EMS personnel, 2. capture & speech recognition system integrated, 3. systematic noise impact study (for speech recognition), 4. two medical drills, 5. medical info extraction from speech

Timeline and Cost

Activities	CY	17	18	19	20
Medical Workflow Studies		[Bar chart showing activity from CY17 to CY19]			
Platform development		[Bar chart showing activity from CY17 to CY19]			
Assessment			[Bar chart showing activity from CY18 to CY19]		
Roadmap development				[Bar chart showing activity in CY19]	
Estimated Budget (\$K)			\$893k	\$751k	

Goals/Milestones

CY17-18 Goals – Medical Workflow and Platform Development

- Focus Groups, Drill #1
- Platform Integration
- Speech Recognition

CY18-19 Goals – Qualitative Assessment and Roadmap Development

- Information Extraction and integration with EHR
- Video Digest
- Drill #1-2, #2
- Workflow Assessment

Comments/Challenges/Issues/Concerns

- Base system and study to date identifies challenges & promising opportunities for further development of robust solutions.

Budget Expenditure to Date

Projected Expenditure: \$1,644,510

Actual Expenditures: \$728,498.59 @31DEC2018

- ❖ Purpose: This research project addresses one common perception of hands free electronic health record data collection for the operational environment
 - ❖ Hands free data entry requires speech recognition, but ...
 - ❖ It is too hard to do in the field with all the noise, movement, and harsh elements

- ❖ Methodology
 - ❖ Decoupled recording and recognition
 - ❖ Robust recording in the field + Just in time speech recognition + perpetual archive
 - ❖ Medical leads the way
 - ❖ Focus group studies with EMS physicians, helicopter transport nurses with prior military medical experiences + EMS drills + systematic quality assessment
 - ❖ Platform for integration
 - ❖ All components are evolvable: Data input devices, compute HW/SW, storage, EHR interface
 - ❖ Resilience and usability rooted in platform design and data lifecycle
 - ❖ Focus on system architecture & data interfaces, instead of individual component performance
 - ❖ Creation of military-oriented training data set for speech and video recognition

- ❖ Expected Outcomes
 - ❖ Medical workflow and requirements
 - ❖ CARD platform prototype hardware and software
 - ❖ Assessment instruments and results

- ❖ Hypothesis(es) / Research Question
 - ❖ The key objective for the solution is two-fold
 - ❖ Complete and resilient documentation (CARD)
 - ❖ Timely data that enhances medical workflow
 - ❖ The key concerns for the solution differ based on roles of care
 - ❖ Point of injury (medics) – record, confirm correct capture
 - ❖ Transport/en route care (nurses) – record further, confirm correct capture, communicate advance stats downstream
 - ❖ Field Hospital (physicians) – review assessment and treatment history since point of injury
 - ❖ Recorded data can be in multiple forms linked to a patient's EHR record
 - ❖ Original audio/video recording
 - ❖ Structured text data via speech recognition and natural language understanding
 - ❖ The former assures complete and resilient documentation. The latter assures useful enhancement to medical care
 - ❖ Speech recognition performance can be persistently improved given
 - ❖ A sizeable archive of recordings for training deep-learning based software
 - ❖ Decoupled “human-audit” and “re-train” process using the archive.
- ❖ Specific Aims / High Level Objectives
 - ❖ Aim 1: Medical Workflow
 - ❖ Identify requirements and operational conditions from physicians and nurses
 - ❖ Study medical rescue standards and deviations in operational environments
 - ❖ Aim 2: Platform
 - ❖ Integrate, end-to-end, cameras, GPU compute devices, storage, to Cerner EHR
 - ❖ Develop and integrate speech recognition, video digest solutions
 - ❖ Explore a multitude of data interface and management options
 - ❖ Aim 3: Assessment
 - ❖ Experienced physician/nurse/EMS drills
 - ❖ Surveys for workflow efficiency
 - ❖ Aim 4: Technology roadmap development

<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Platform</p>	<p>Camera + sensors GPU platforms EHR</p>
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Software</p>	<p>NLP-based speech and video information extraction</p>
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Workflow & Assessment</p>	<p>Focus Groups Instrument Drills</p>

TACTICAL COMBAT CASUALTY CARE (TCCC) CARD

BATTLE ROSTER #:

EVAC: Urgent Priority Routine

NAME (Last, First): _____ LAST 4: _____

GENDER: M F DATE (DD-MMM-YY): _____ TIME: _____

SERVICE: _____ UNIT: _____ ALLERGIES: _____

Mechanism of Injury: (X all that apply)

Artillery Blunt Burn Fall Grenade GSW IED

Landmine MVC RPG Other: _____

Injury: (Mark injuries with an X)

TQ: R Arm TYPE: _____ TIME: _____

TQ: L Arm TYPE: _____ TIME: _____

Focus Group & Drills

❖ Objectives:

- ❖ Focus groups with EMS, transport nurses, ED physicians → Needs & requirements
- ❖ Battlefield rescue drills in sim center → A/V recording for S/W dev & testing

❖ Approach:

- ❖ 6 focus groups, total 26 participants
- ❖ 2 drills: 3 trauma types x 3 stages x 2 medics x 1 soldier, scripted speech

❖ Findings:

- ❖ Focus group: multiple speakers, multiple patients, chaotic conditions, need of on/off control
- ❖ Drills: 4 hrs recording, multiple noise levels

Platform Design

❖ Objectives

- ❖ Not to reinvent the wheels
- ❖ Identify unique challenges
- ❖ Propose and validate near- and long-term improvement strategies

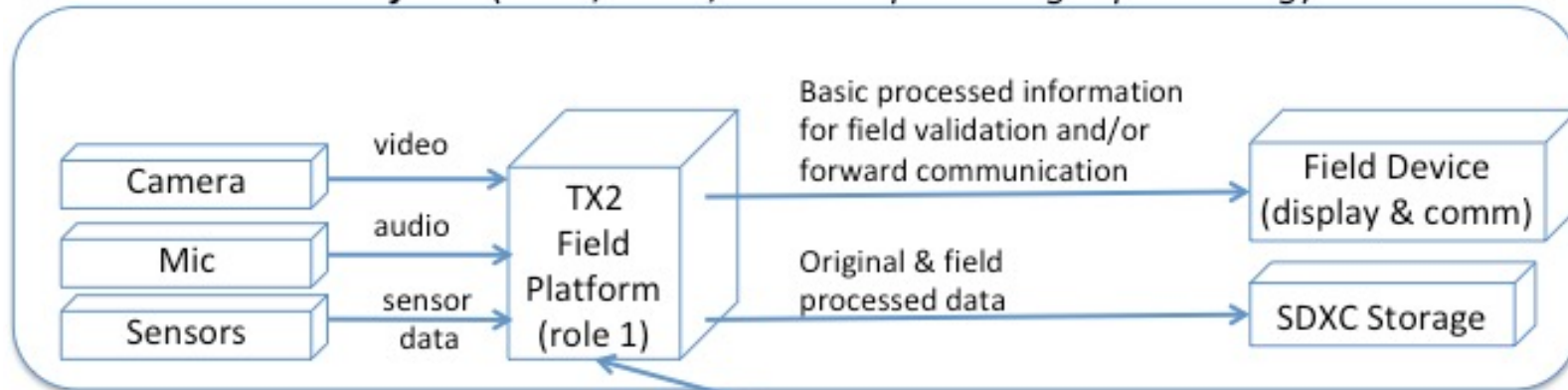
❖ Principles

- ❖ Examine holistically, end-to-end
- ❖ Analyze incisively, modular
- ❖ Long term persistent enhancement
- ❖ Open integration

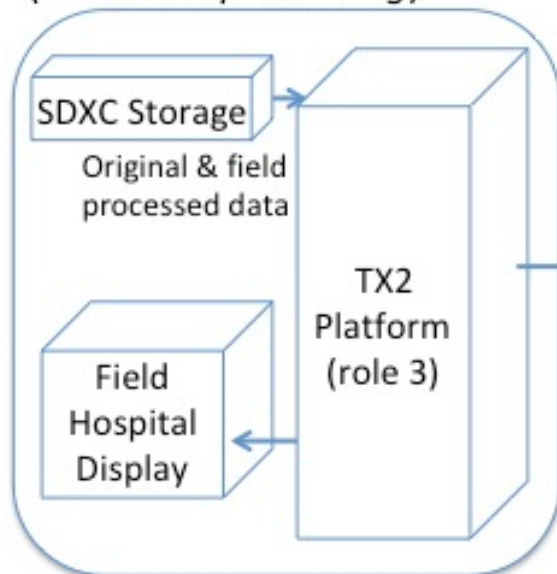
❖ Approach

- ❖ Systematic, modular performance analysis (varied noise, S/W and H/W components)
- ❖ Open source, modular S/W (instead of end-to-end black-box)

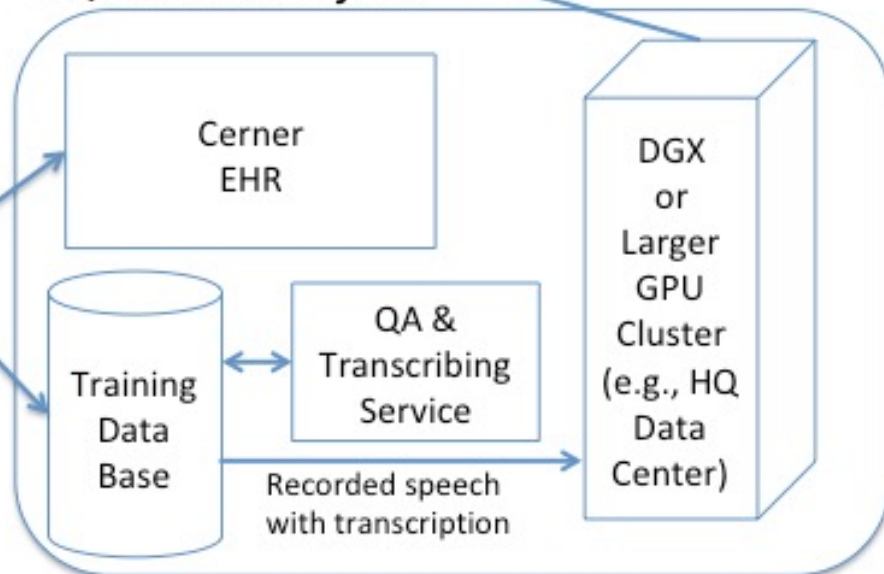
Role 1 Field Platform (video/audio/sensor capture + light processing)

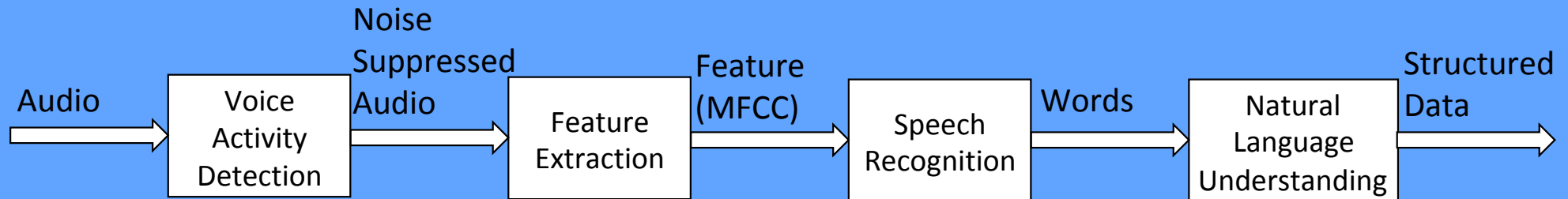


Role 3 Field Hospital Platform (enhanced processing)



HQ Backend Platform

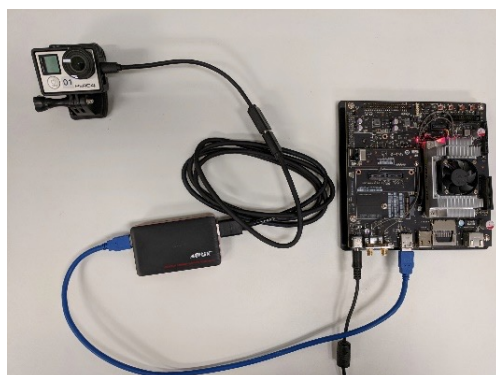




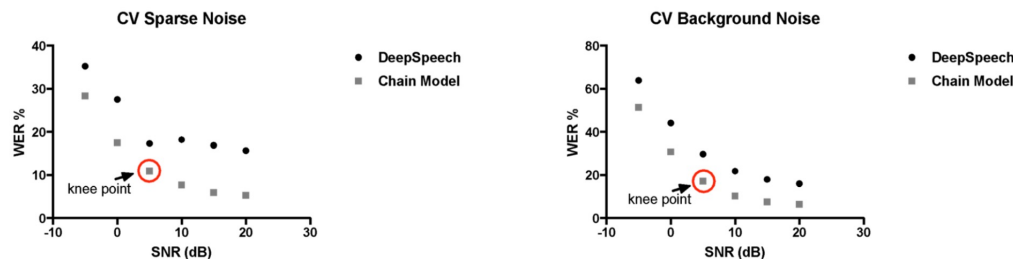
- Focus – understand battlefield-specific challenges & design for sustainable modular upgrade:
 - ❑ **Noise.** Establish specific relationship btw. noise and speech recognition performance
 - ❑ **Language.** Establish specific vocabulary used by (military) medical operations
 - ❑ **Modularity.** Establish framework for continued modular upgrades

- Methodology
 - ❑ Hardware (binaural microphones) and software noise-reduction
 - ❑ Hybrid deep learning (Kaldi) instead of end-to-end deep learning (DeepSpeech)
 - ❑ Pre-processing and parallelization for multi-speaker audio
 - ❑ Speech enhancement, medical language model training (NLP)
 - ❑ Post-processing for EHR information extraction

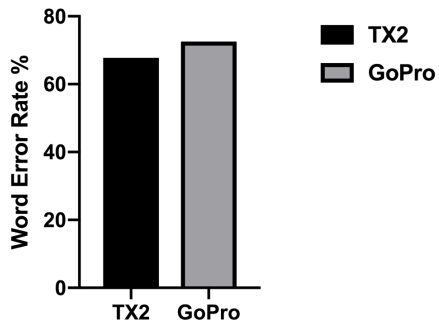
- ❖ Speech recognition
 - ❖ on-medical system complete, working on speech recognition performance benchmarking, multiple enhancements, information extraction
 - ❖ Metrics: general word error rate, medical word error rate
- ❖ Platform integration
 - ❖ Completed – Camera + video capture + TX2 GPU board + software + battery



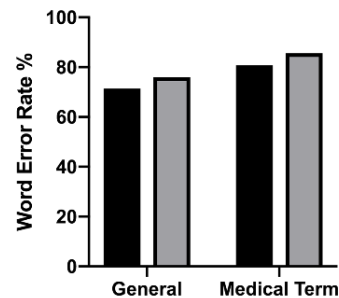
- ❖ Cerner integration
 - ❖ Working with Cerner to address challenges in uploading documents to sandbox



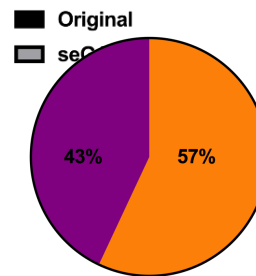
- ❖ Simulation Center Drill #1, #1-2
 - ❖ 4 hours video/audio recording in three simulated settings
 - ❖ GoPro and NVIDIA TX2
 - ❖ Key findings – microphone matters, language model & multiple speakers are key!



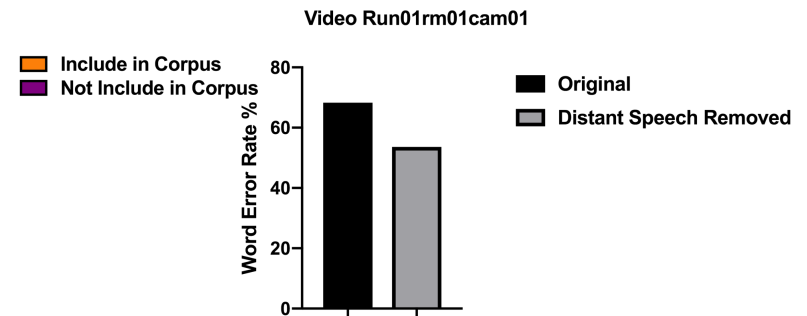
	TX2	GoPro
WER	67.74	72.58



	Original	seGAN
General	71.41	75.96
Medical Term	80.80	85.60

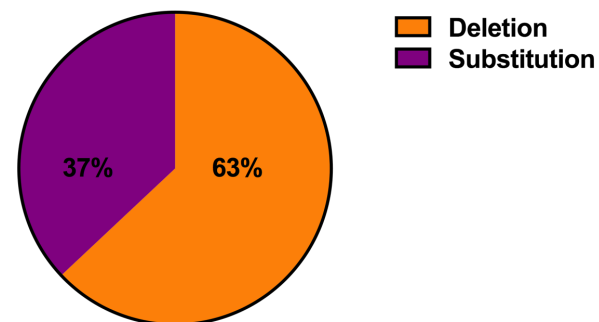


Total Medical Terms



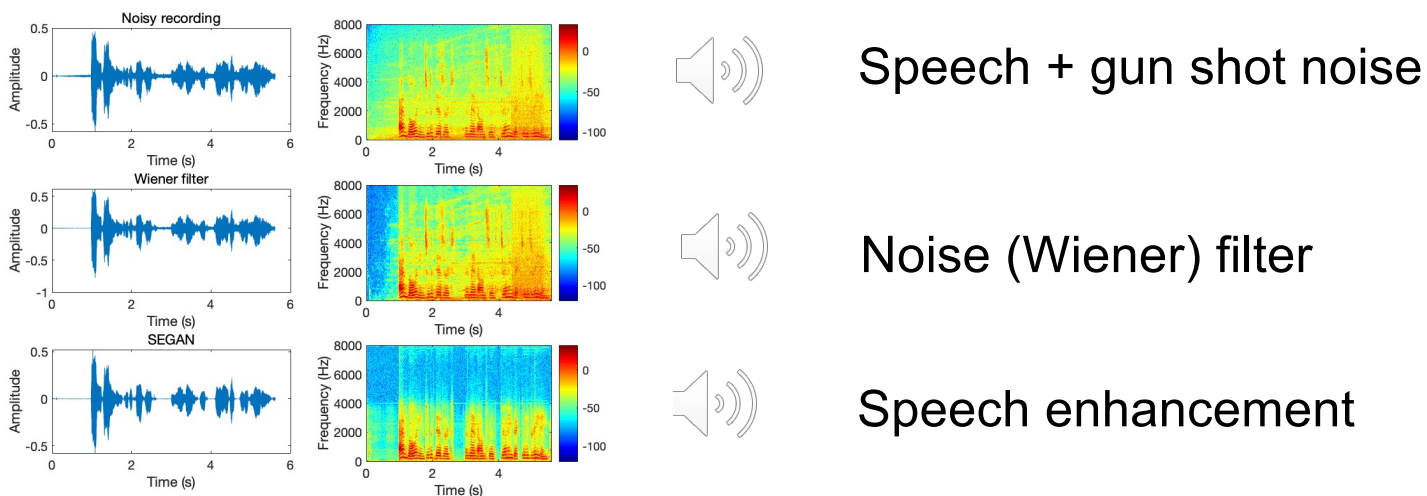
Category	WER
Original	68.27
Distant Speech Removed	53.63

- Discerning types of speech recognition errors
 - » Deletion – caused by noise or “distant” speaker heard
 - » Substitution – caused by insufficient vocabulary
 - » Solutions under way:
 - Training S/W with specific medical terms needed
 - Speech enhancement
 - Parallel processing of multiple medic recordings



Medical Term WER

- Effects of noise filtering & speech enhancement



❖ COST

- ❖ What is the Money(\$) spent to date? \$728,498.59
- ❖ Is burn rate on schedule? Yes
- ❖ Have you spent too much or too little; and why?
This project is financially on track.

❖ SCHEDULE

- ❖ Is the project on schedule, behind schedule, what delays?
This project is on schedule.

❖ PERFORMANCE

- ❖ What tasks have you successfully completed on your SOW?
 - ❖ Task 1.1: Focus Group/Interviews
 - ❖ Task 1.2: First EMS Drill; subsequent Drill 1-2 (12/4/18)
 - ❖ Task 2.1: Platform Integration
- ❖ What tasks are at risk ? There are no tasks at risk.

GANTT Chart



	Y1Q2	Y1Q3	Y1Q4	Y2Q1
Specific Aim I- Medical Workflow Studies				
Task 1.3: Qualitative Assessment (80%)				
Quality assessment instrument design				
Survey execution and analysis				
Specific Aim 2- Platform Development				
Task 2.1: Platform Integration (100%)				
Acquisition of components				
Platform and data interface integration				
Cerner interface validation				
Speech and video software integration				
Task 2.2: Speech Recognition & Video Digest				
Basic speech recognition development (100%)				
Training data preparation & training (100%)				
Video digest (40%)				
Cerner integration (40%)				

GANTT Chart



	Y1Q2	Y1Q3	Y1Q4	Y2Q1
Specific Aim 3 – Assessment				
Task 3.1: Technology Assessment				
Platform benchmarking (60%)				
Drill #2 design and preparation (40%)				
Task 3.2: Workflow and EHR Assessment				
Qualitative Survey of workflow quality (40%)				

Specific Aim 1 – Medical Workflow Studies

Task 1.3: Qualitative Assessment

Quality assessment instrument design (months 7-12) – 80% Completed -This task will define quality goals and instrument needed to assess the medical effectiveness of CARD. To date, we have decided to base the design on the DD 1380 TCCC card. Remaining work focuses on determination of information priority and CARD integration requirements. We have confirmed use of the TCCC CARD format for the EHR report. Based on that format, we were able to define more specific metrics for accessing speech recognition performance for specific medical information. This leads to current work on information extraction software based on the speech recognition output. Completion of design of assessment instrument and assessment procedure in Q5.

Survey Execution and analysis (months 13-18) – Conduct drill #2 as additional development drill with preliminary survey.

Specific Aim 2 – Platform Development

Task 2.2 – Speech Recognition & Video Digest

Video digest (months 10-15) – 40% Completed - In addition to the already completed basic video digest function linking video footage based on timestamp of recognized speech, in Q5, we will explore additional, more challenging video analysis such as facial identification of known persons.

Cerner integration (months 10-15) – 40% Completed - After contacting Cerner FHIR, it was confirmed that Cerner FHIR does not support the interface for external software to upload data into the patient record. We will continue to work with Cerner to identify road map for integration and set up a local server to display data records from the proposed project work.

Specific Aim 3 - Assessment

- Task 3.1: Technology Assessment
- **Platform benchmarking** (months 10-12) – 60% Completed - Platform benchmarking will continue into Q5 with new metrics and assessment methods.
- **Drill #2 design and preparation** (months 10-15) – 40% Completed -In response to new assessment plan (Tasks 1.3, 3.1), drill #1-2 is re-positioned as a developmental drill to acquire recordings of medical dialogues consistent with the assessment standard. An additional drill (drill #2) is planned beyond drill #1-2 for full assessment.
- **Drill #2 execution** (month 18) – Based on new plan, drill #1-2 was for developmental purpose and may take place earlier than month 18. In addition, drill #2 will be planned for full assessment.

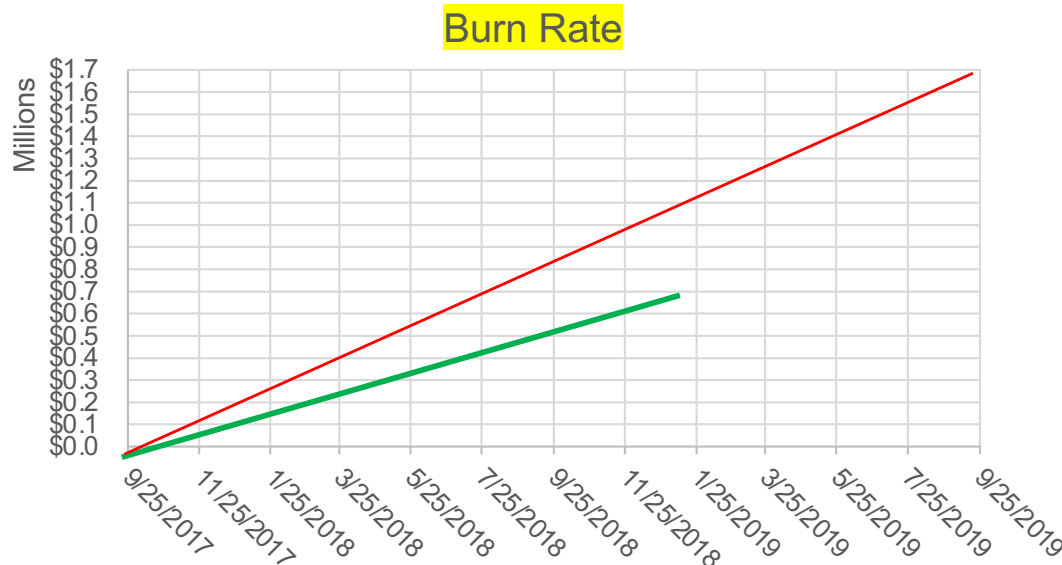
- Task 3.2: Workflow and EHR Assessment
- **Qualitative Study and workflow quality** (months 13-18) – 40% Completed

Specific Aim 4 – Roadmap Development

- **Platform usability and gaps assessment** (months 19-24) .
- Roadmap development (months 19-24) .

-

- ❖ Budget Expenditure to-date:
 - ❖ Projected Expenditure: \$1,644,510 Actual Expenditure: \$728,498.59 @ 12/31/19
 - ❖ Anticipated adjustments on remaining periods - None
 - ❖ Due to delays in hiring a graduate student with a much-desired expertise in speech recognition, as well as a similar challenge in identifying a post-doc with a fitting expertise, we have delayed spending in salary line. The graduate student started in August 2018 and immediately made significant contributions. In addition, since June 2018, we remedied the situation by hiring a post-doctoral researcher on a part time basis (due to her limited available time) to the project team leading to scientific publication efforts.
- ❖ Speak to expected vs. actual burn rates
 - ❖ Are you spending money too fast or too slow, and why? This study is financially on track.



Contracting Overview



- ❖ HIGHLIGHT tasks that have deviated from original proposal or SOW.
 - ❖ At this point, there is no deviation but we do see opportunities to deepen the original intended SOW

Deliverables

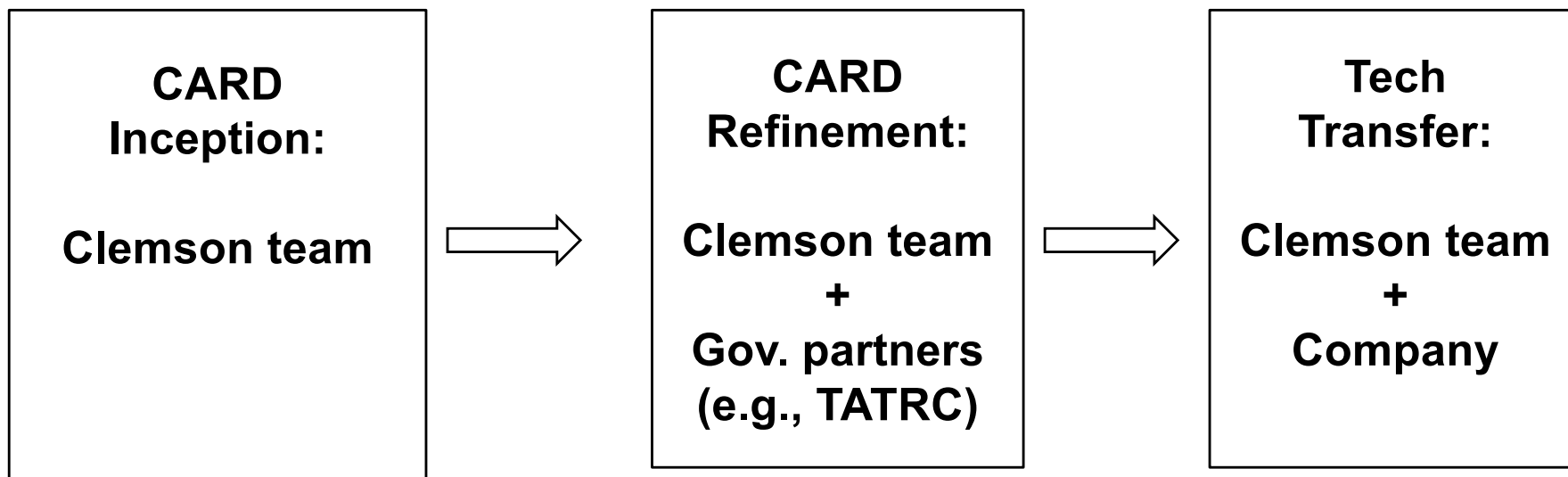


Deliverables & Data Requested	Details/Format	Data Rights	Delivery Schedule
Quarter #2 Technical Progress Report and Quad Chart	A brief description of overall progress to date for the reporting period 3/26/18-6/25/18	Full access	3/24/18
Quarter #3 Technical Progress Report and Quad Chart	A brief description of overall progress to date for the reporting period 6/26/18-9/25/18	Full Access	6/24/18
Annual Progress Report and Quad Chart	A description of overall annual progress to date for the reporting period 9/26/18-9/25/18	Full access	9/25/18
REVISED Annual Progress Report and Quad Chart	A more detailed description of overall annual progress to date for the reporting period 9/26/17-9/25/18	Full access	11/20/18
Focus Groups	Griffin, S., Gimbel, R., Wang, K., Edwards, K., Gerard, W., Shelton, S., Hawsey, K. Complete and Resilient Documentation (CARD) for Operational Medical Environments: Knowledge Gained from Focus Groups Regarding the Use of GoPro® Equipment in Military and Civilian Medical Environments, 2018 MHSRS, August, 2018	Full access	Aug, 2018

Anticipated Impact as an Outcome of Research



- ❖ The research expects significant impacts in:
 - ❖ Commercial, Civilian and Military
 - Commercial development in key technologies
 - Synergistic use in civilian EMS, homeland security and disaster recovery
 - New companies or consortium of companies on standard based integrated platform, from variety of data capture devices, mobile app, to cloud solutions
 - Integration with telemedicine, enhance the care and survivability of service members
 - Establish understanding of battlefield readiness of state-of-the-art technology
 - Establish foundations for long-term success for DoD speech recognition competencies
 - ❖ Academic
 - ❖ Speech recognition, natural language understanding, and video uses in battlefield settings is a rich area for academic research
 - ❖ Large scale, distributed data collection (speech, video, sensors) and aggregation into individual patient data is a rich area of research



Military & academic research in speech, NLU & clinical info extraction & use in operational environments

Civilian applications for emergency medicine & homeland security

- ❖ As it completes the Q5 (Y2Q1), the project is on track on all proposed tasks.
- ❖ The anticipated outcomes and their status/outlook are as follows:
 - ❖ Focus group study – study completed, publication in progress
 - ❖ Platform development and concept demonstration – basic speech recognition completed, performance study in progress, natural language understanding and video digest to begin
 - ❖ (2nd) simulation drill & quality assessment
- ❖ The project's aim is not a point product. Instead, its aims are on:
 - ❖ Confirmation of the CARD concept's clinical opportunities and challenges
 - ❖ Development of the platform's architecture and interfaces
 - ❖ Demonstrating the extensible nature of the platform components
 - ❖ Assessing the CARD design's impact to the medical workflow

BACK-UP SLIDES



BACK-UP SLIDES



BACK-UP SLIDES



	Total n	Male n (%)	Female n (%)		Experience <10 years n (%)	Experience 10-20 years n (%)	Experience >20 years n (%)	Experience Unknown N (%)
EMS	6	5(83%)	1 (17%)		1 (17%)	1 (17%)	3 (50%)	1 (17%)
Transport Nurses	11	8 (73%)	3 (27%)		1 (9%)	2 (18%)	8(73%)	0 (0%)
ED Physicians	9	8 (89%)	1 (11%)		3 (33%)	2 (22%)	4 (44%)	0 (0%)
Total	26	21(81%)	5 (19%)		5 (19%)	5 (19%)	15 (58%)	1 (4%)

Focus Group Conclusions

Informed protocol needs for collecting and transmitting data

Suggests information (notes from audio & video) could be useful at multiple phases of care

Audio recognition sophistication for chaotic communication



- ❖ Palmetto Health System IRB Approval – 10/25/17
- ❖ Clemson University IRB Approval – 10/26/17
- ❖ HRPO Military Secondary Review Approval – 11/17/17

❖ Poster presentations/Publications/Articles

- ❖ Griffin, S., Gimbel, R., Wang, K., Edwards, K., Gerard, W., Shelton, S., Hawsey, K. Complete and Resilient Documentation (CARD) for Operational Medical Environments: Knowledge Gained from Focus Groups Regarding the Use of GoPro® Equipment in Military and Civilian Medical Environments, 2018 MHSRS, August, 2018

- US Army News: <https://www.armytimes.com/news/2017/12/18/device-would-help-combat-medics-document-treatment-in-the-field/>
-
- Post and Courier: https://www.postandcourier.com/features/clemson-researchers-envision-an-improved-way-to-document-battlefield-injuries/article_24464cf6-dec6-11e7-b159-4f4c3a15847e.html
-
- Clemson World: <https://clemson.world/witnesstowar/>
-
- Clemson News: <http://newsstand.clemson.edu/mediarelations/injured-troops-could-receive-better-care-and-benefits-with-new-research/>

❖ Intended future Publications / Presentations

- ❖ A formal report with recommendations will be generated to advise the CARD technology development team and will be delivered to USAMRMC

- ❖ A scientific manuscript regarding the focus group design, recruitment, interviews, transcription, analysis and operation recommendations will also be crafted.



Complete And Resilient Documentation
for Operational Medical Environments -
Challenges and Opportunities for Automatic
Speech Recognition

Kuang-Ching (KC) Wang

Professor, ECE
Associate Director of Research, the Watt Center
Clemson University

The Challenge



“Today, there is about 0% point-of-injury patient documentation of US combat rescues.”
source: unnamed U.S. military official

The Dilemma

- Significant R&D to-date in hands-free documentation technologies with limited success.
 - Speech recognition¹
 - Bone conduction²
 - Sensors
 - Thriving success in commercial speech recognition technologies and applications.
 - “Conventional” model-based methods
 - “End-to-end” deep-neural-network based methods
 - Apple, Amazon, Google, Microsoft, Baidu, Dragon (Nuance)
1. Speech Understanding Workshop, 1975.
 2. “Osophone” hearing aid, 1923.

To Do, or Not To Do (Speech Recognition)

- Let's do it
 - Concern of errors
 - https://github.com/syhw/we_are_we
 - Up to 50% error under certain conditions, e.g., noise, multiple speakers
 - How to decide if technology is ready to be fielded
- Let's not do it
 - Other hands-free documentation methods?
 - Gesture recognition (Vanderbilt Univ.)
 - Status quo is not acceptable
 - Tremendous veteran testimonials of hardship due to lack of documentation

Untangle the Needs

- Documentation

- Ground truths for injury

- Audio and Video
 - The wounded
 - The medics



**Complete And Resilient
Documentation
"CARD"**

- Speech recognition

- Transform ground truth into digital form for software applications

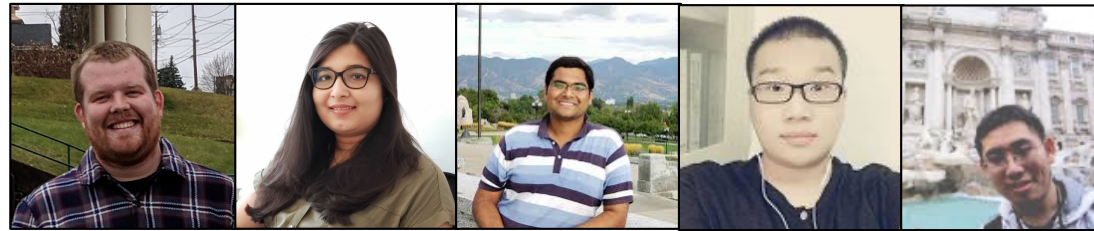
- How urgent?

- Near real time
 - Advance notification to field hospitals
 - Telemedicine
- Post rescue
 - EHR records
 - Recovery care and benefit claims

The CARD Team



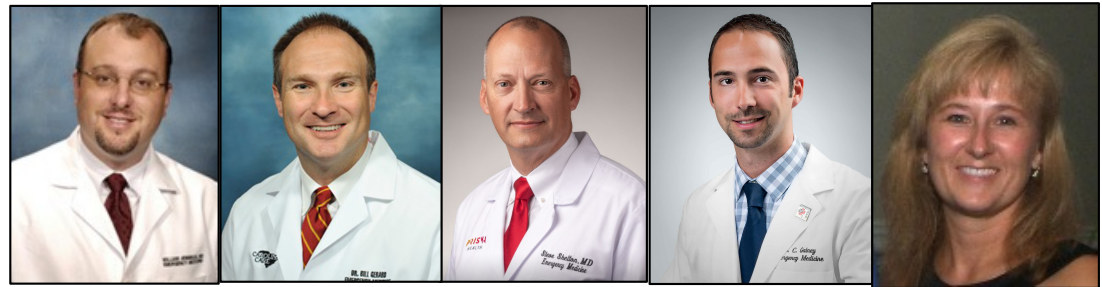
ECE (networking,
machine learning)



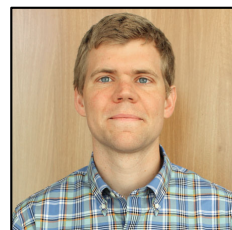
ECE (platform, speech recognition, experimentation)



PHS (military medical operations,
qualitative studies)



Palmetto Health Emergency Medicine
(study groups, medical simulation)



BDSI, ICAR, Watt Center, Computer Science
(AI, Natural Language Processing)

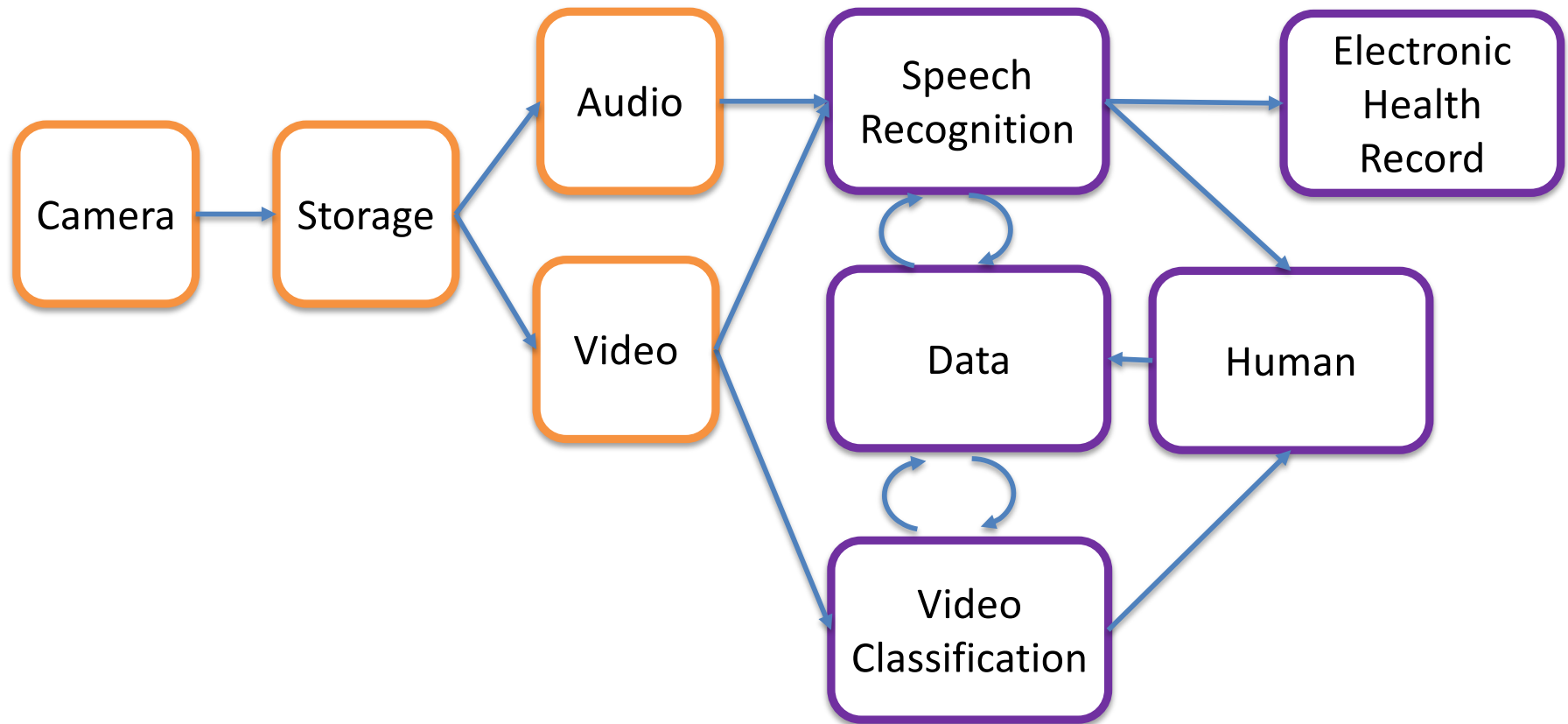
Focus Groups with EM Physicians and Nurses

Conclusions

- Informed protocol needs for collecting and transmitting data
- Suggests information (notes from audio & video) could be useful at multiple phases of care
- Audio recognition sophistication for chaotic communication

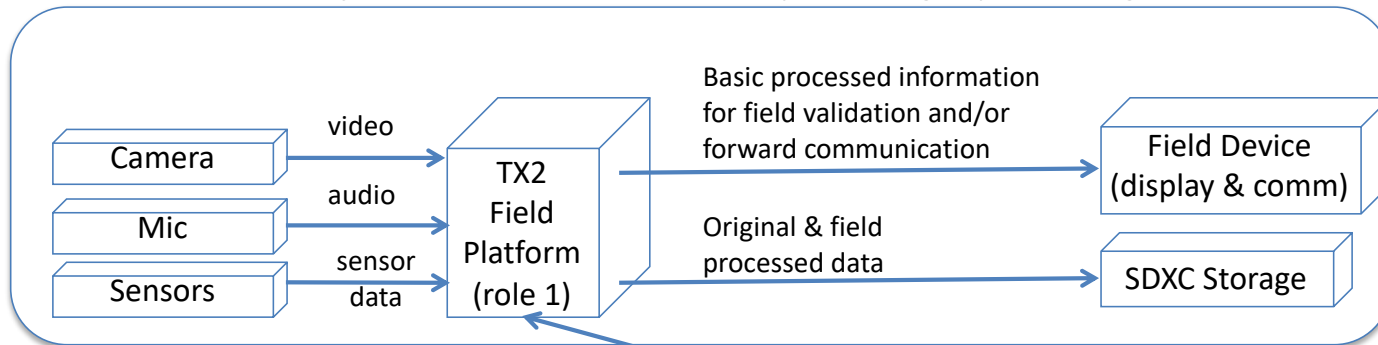


Decompose the Process



A System's Approach

Role 1 Field Platform (video/audio/sensor capture + light processing)



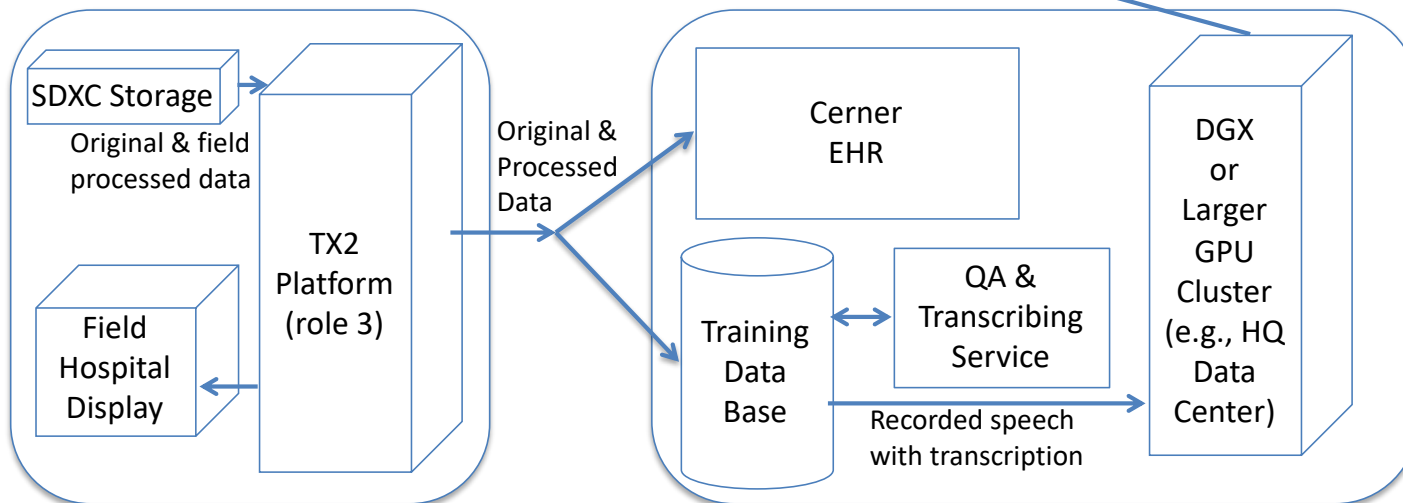
US military roles of care:

- Role 1 – first responder
- Role 2 – forward resuscitative care
- Role 3 – theatre hospitalization

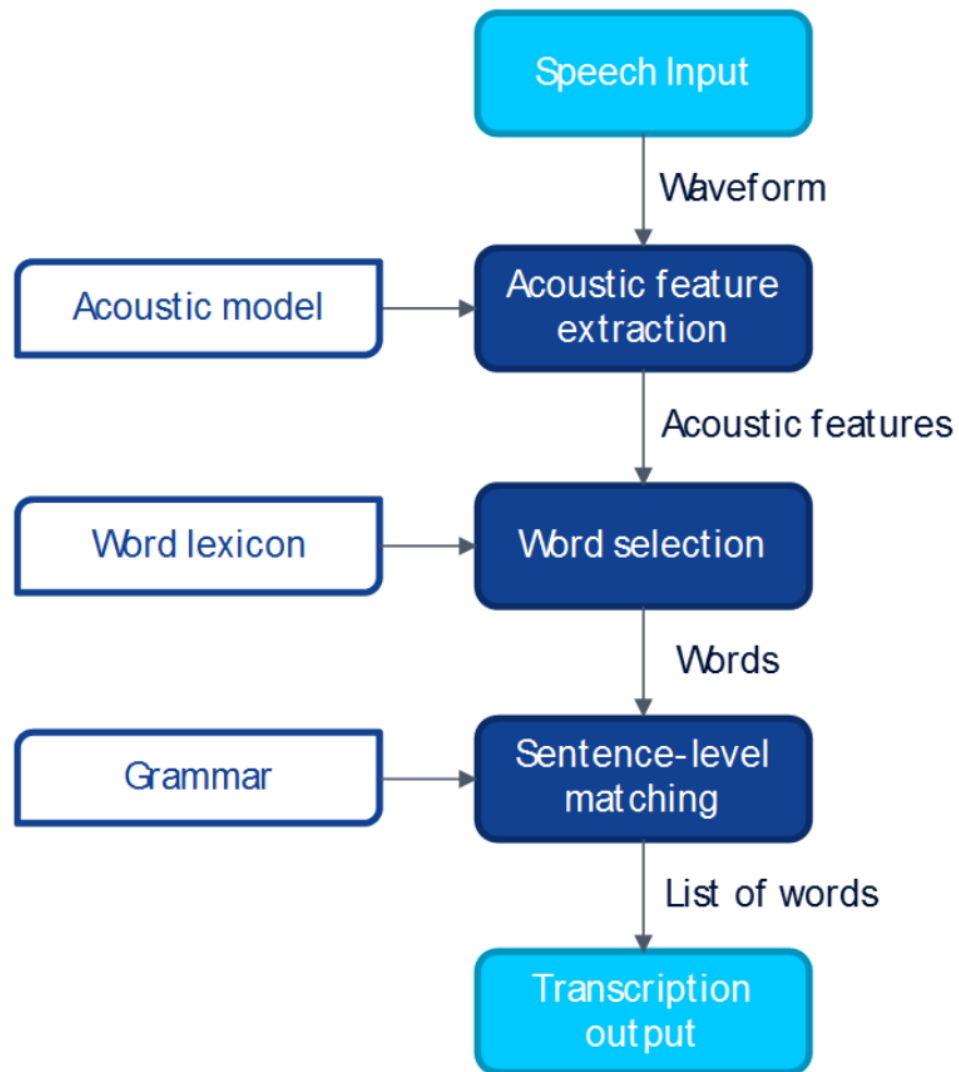
Continued upgrade
via software live updates

Role 3 Field Hospital Platform (enhanced processing)

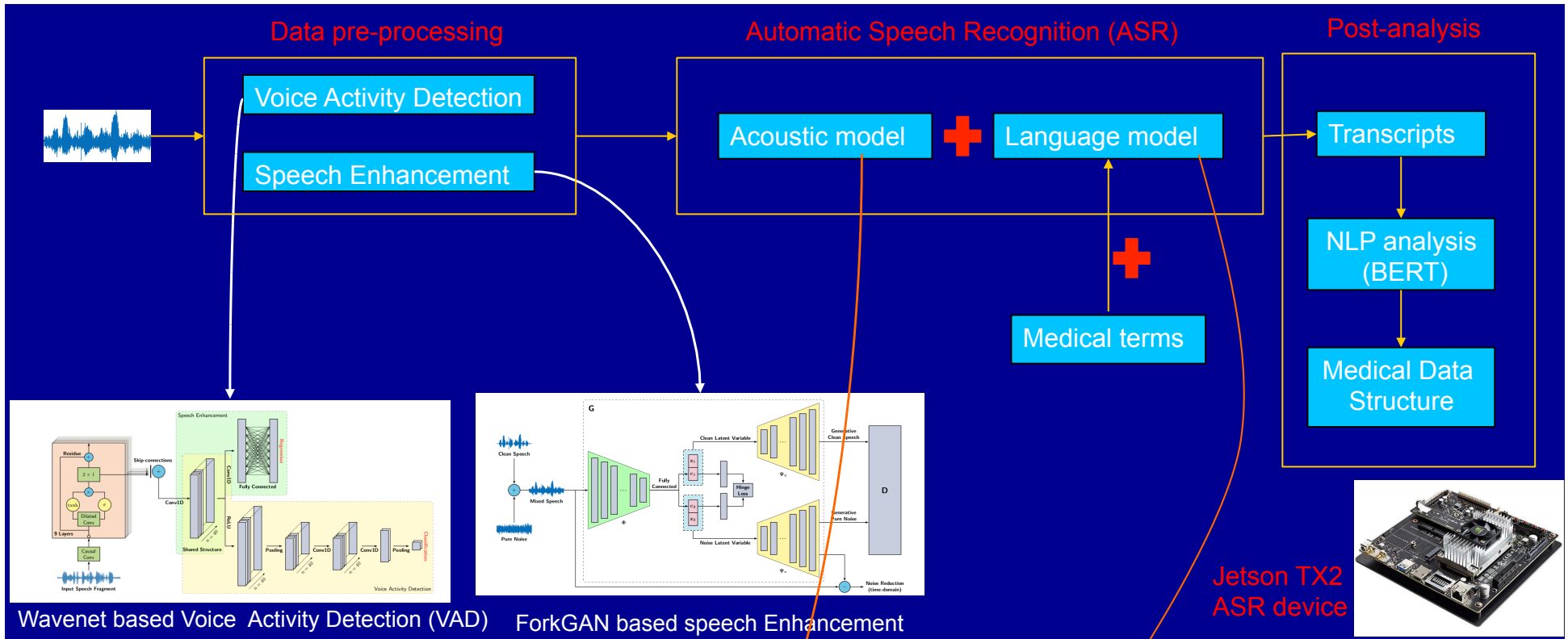
HQ Backend Platform



Speech Recognition: Model-based Workflow

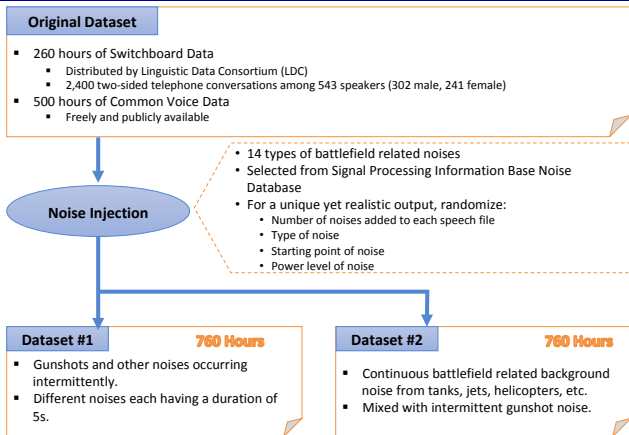


Speech Recognition: Hybrid, Model-based – Kaldi



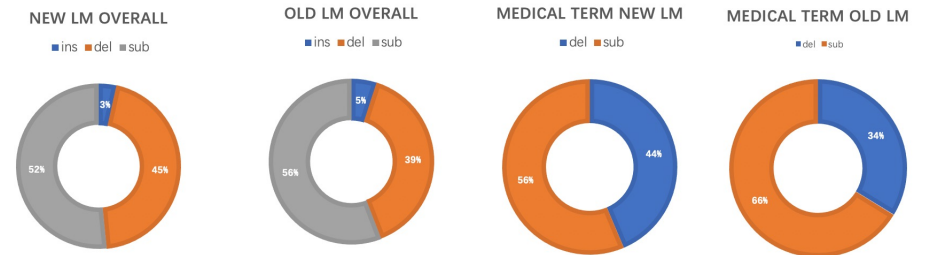
Wavenet based Voice Activity Detection (VAD) ForkGAN based speech Enhancement

Jetson TX2 ASR device



Test data: 7 drill recording (2994 words)

Evaluation	WER	Medical WER
OLD LM	69.77%	81.19%
New LM	64.76%	52.0%



Data Sets

- Speech data sets
 - Switch Board
 - Common Voice
 - Libris Speech
 - Medical drills
- Noise augmentation
 - 14 types of battlefield noises
 - Intermittent and continuous noises

Errors, Lots of Errors

- Show WERs, pie charts based on error types

What Can We Do

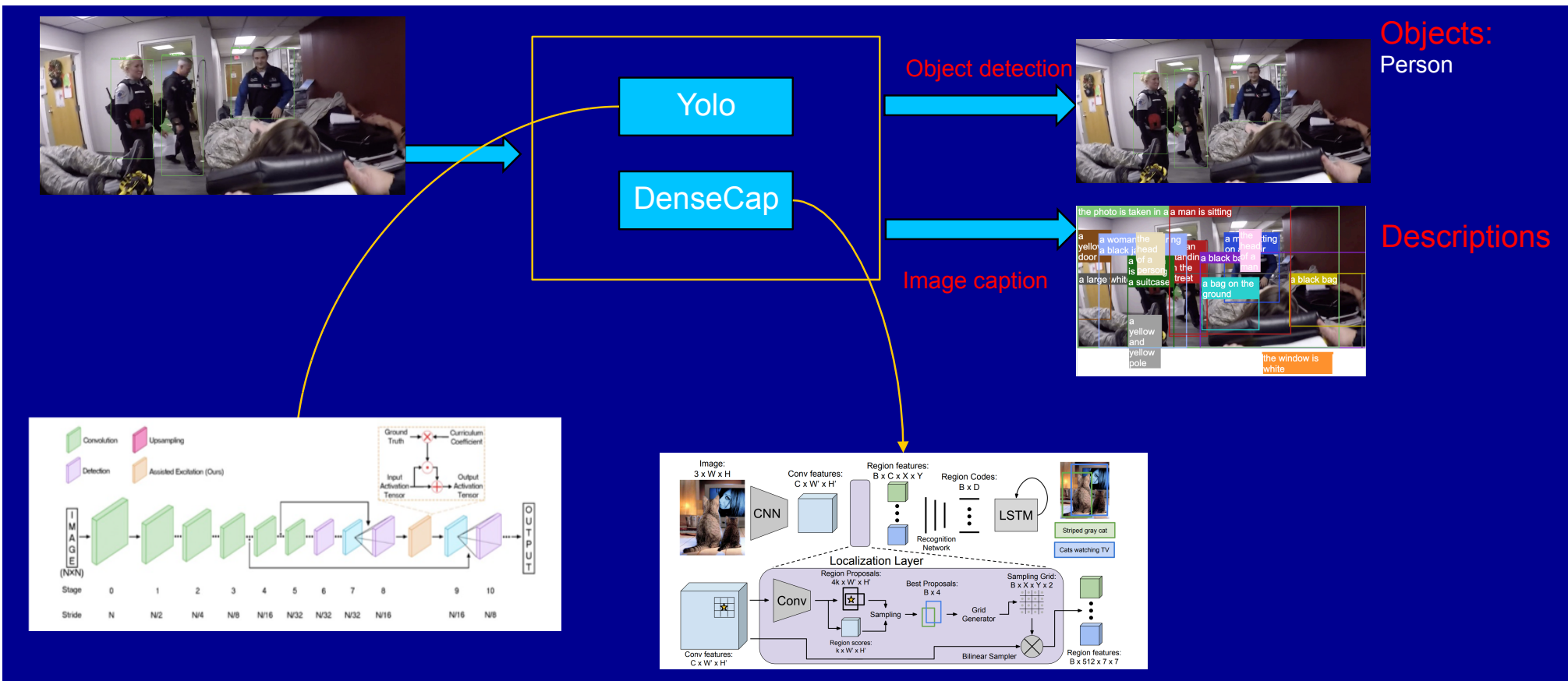
- Speech Enhancement
- Language Model Enhancement
- Post-hoc Correction

More about the Errors

- Examples of errors and discussion of suitable methods for improvement

Post-hoc Correction

Video Analysis Workflow – Preliminary study



Conclusion