Eye Motion Tracking for Desktop-Based Medical Image Interpretation Training

Chanler Cantor & William Marx, Ph. D.
Internal Research and Development,
Intuitive Research and Technology Corporation
Huntsville, AL
chanler.cantor@irtc-hq.com;
william.marx@irtc-hq.com

Junjian Huang, M.D. & Andrew Smith, M.D. Ph.D.
Department of Radiology, Heersink School of
Medicine, University of Alabama at Birmingham
Birmingham, AL
junjianhuang@uabmc.edu;
andrewdennissmith@uabmc.edu

Matthew Schultz
Daft Resolution Studios
matthew.r.l.schultz@gmail.com

ABSTRACT

The USAF Pilot Training Next (PTN) initiative demonstrated the value of Eye Motion Tracking (EMT) in augmenting learning by guiding the trainee to know where and when to look; how to search, detect, and/or recognize patterns; and how long to gaze. The purpose of this research is to demonstrate the benefits of obtaining real-time feedback of eye direction and gaze duration to trainees learning about non-military medical imaging use-cases. This real-time indicator can allow the trainer to adapt verbal queueing of the trainee to improve knowledge transfer, and to increase the confidence of the trainer and the competency of the trainee. EMT is popular in professional online gaming, providing amateur e-sport fans a method of watching precisely where professional players' eyes are looking as they play complex games with data-intensive user interfaces. The authors are evaluating EMT systems and developing protocols to improve training in interpretation of medical images. Our initial experiment data set included bone radiographs, digital subtraction angiograms, and computed tomography (CT) images. Preliminary results indicated that expert viewers were able to use EMT to successfully guide novice readers through search and gaze protocol patterns. Although significant research has been conducted on the benefits and applicability of EMT for medical image interpretation training, there has not been widespread adoption of EMT into the training curricula. This paper presents a pilot/prototype project to create an EMT-augmented radiology training system. We provide details associated with our prototype system development, rationale and selection of medical studies, training protocols, and preliminary results. We conclude with a summary of initial lessons learned, and immediate future direction for both technical system development and clinical evaluation within the radiology training community.

ABOUT THE AUTHORS

Ms. Chanler Cantor is an Area Leader at Intuitive Research and Technology Corporation. She is responsible for managing the Internal Research and Technology portfolio and leads multiple projects associated with Artificial Intelligence and Machine Learning, Big Data Analytics, complex visualization, and medical imagery visualization. She and her team develop technology demonstrators and prototypes which can be showcased to customers to promote awareness and early evaluation and adoption of technologies. She has been awarded six patents related to big data visualization and other emerging technologies. She received her BS in Electrical Engineering from The University of Alabama and is currently pursuing her MS in Systems Engineering from Johns Hopkins University.

Dr. William Marx is the Senior Vice President and Chief Technology Officer of Intuitive Research and Technology Corporation in Huntsville, Alabama. He is responsible for planning, managing, and executing research and development programs aligned with the technology priorities of the US military and commercial customers. His experience base and technical portfolio include advanced visualization systems, Big Data analytics, Artificial Intelligence and Machine Learning, Knowledge Based Systems, missile system design, multi-disciplinary design optimization, missile guidance and control, analysis of exo-atmospheric kill vehicles, supersonic aircraft design, and ground and space robotic system design. Dr. Marx received his PhD and MS in Aerospace Engineering from the Georgia Institute of Technology and his BS in Aerospace Engineering with a minor in Mathematics from Embry-Riddle Aeronautical University. He was a NASA Langley GSRP Fellow.

Dr. Junjian Huang is physician scientist at the University of Alabama at Birmingham where he serves as the Director of Medical Student Research for the Department of Radiology as well as Director of Clinical Research for the Division of Vascular and Interventional Radiology. He is responsible for guiding research aims within his division in addition to educating Radiology residents on how to develop procedural expertise. Dr. Huang received his MD at the Medical College of Georgia, and is Alumnus of Mayo Clinic Surgery, University of Pennsylvania Radiology, and Harvard/MGH Interventional Radiology.

Dr. Andrew Smith is a physician scientist at the University of Alabama at Birmingham where he is professor and Vice Chair of Innovation, Co-Director of Augmented Intelligence in Radiology, Director of the Human Imaging Core Lab (HUMI), and founder of two start-up companies related to artificial intelligence in medical imaging. His role is to serve as a consultant for this project as well as a mentor to Dr. Huang. Dr. Smith received his MD and PhD in molecular and cellular biology at the Medical College of Ohio and is alumnus of Cleveland Clinic.

Mr. Matthew "Gus" Schultz is a Software Engineer at Daft Resolution Studios. He is responsible for evaluating existing and developmental software solutions for viability, usability and maintainability, development of software solutions, as well as providing analysis and recommendations to minimize risk and maximize product effectiveness. His experience base and technical portfolio include web development, game development, video content creation, Patriot Missile System maintenance, and technical training development and conduct. Mr. Schultz received his BS in Game and Simulation Programming from DeVry University.

INTRODUCTION

Eye tracking detects and measures the point of gaze and the motion of the eyes as they search and stare at objects in 3D space or at a 2D plane. In recent years, eye tracking technology has made significant improvements in accuracy, stability, and sampling rates (Fu et al., 2016). This technology continues to mature, providing an opportunity for research, evaluation, and adoption in multiple fields. Beyond basic research, other applications are advancing in areas such as visual content analysis, design evaluation, usability and consumer behavior research, driving, gaming, and medical research (Valliappan et al., 2020).

Training in diagnostic and interventional radiology requires the development and refinement of pattern recognition skills. Pattern recognition can improve with repeated training and a validated measure for acquisition of pattern recognition competency. Modern eye tracking systems can be used in the competency assessment. Furthermore, they may be used to accelerate the training and learning process when coupled with verbal ques from expert attending radiologists or software cues from the training software itself. Eye tracking software has been shown to be a strong method to assess radiology resident search and pattern recognition skills. However, eye tracking technology has yet to be implemented broadly in the education of radiology trainees.

There are certainly parallels in the United States military training domain. For technology to be widely adopted across the defense enterprise, it needs to be woven into bureaucratic structures and processes, which are inherently sticky (Blanken et al., 2018). The established acquisition cycles, milestones, and processes used to buy new systems to defend our nation were not designed to provide rapid, efficient adoption of technology. This is true for military systems ranging from aircraft and missiles to training systems. The medical training community faces similar challenges for technology adoption. Technology benefits must be objectively proven and validated before systems wide adoption will occur.

The United States continues to face a growing shortage of physicians due to increasing workloads, burnout, and other causes. This long-term trend, coupled with the pandemic and increasing image demand has led to a highly competitive radiologist employment market at a time when medical practices find themselves short-staffed, overworked, and in need of relief (Walter, 2021). The supply and demand issues that have plagued our nation for the past two years are indeed affecting the radiology industry as well. Many radiologists postponed retirement in 2020 due to the pandemic, financial uncertainty, and inability to travel. As some elements of our society have begun their return to pre-pandemic conditions, retirements have taken place, further stressing the supply of specialists available to read exams. Near-term, this causes salary escalation and stresses the education and training systems to provide increasing numbers of qualified radiologists to meet the continuing surge in demand. The aerospace workforce faces similar problems – as our military systems continue to become more complex and data-centric, the supply of computer scientists, software engineers, data scientists, and network architects struggles to maintain pace.

Fortunately, there are technology-based methods that can begin to mitigate growing demands placed on our medical imaging and training communities:

- The development and adoption of technology-based workflow process aids can help processing bottlenecks in our healthcare system. Visweswaran et al., (2021) discuss an application of EMT to collect training data for a machine learning-based clinical decision support tool. The idea is to predict which patient data are relevant for a clinical task. This can allow the system to highlight or focus staff attention on a subset of data to improve processing time.
- Technology can also be used to accelerate the imaging workflow process itself. NVIDIA's Clara Imaging suite is being developed to provide AI-assisted workflows with applications including labeling, modeling, and annotation. Development and use of Clara and similar applications has the potential to help radiologists and imaging centers increase throughput for reading and diagnosis.
- This paper addresses the problem of training medical image professionals to enhance their diagnostic skills by proposing an integrated EMT system designed to increase both radiology trainer and trainee confidence in the trainee's reading ability. This system can also provide a more thorough training mechanism where training can be tailored to the trainee in real time. This increase in confidence and more focused training can be used to reduce the training time needed to transition radiology residents from novice to expert readers. A reduction in time for residents to achieve scan reading competency can begin to mitigate the increasing demands on the imaging community.

EMT FOR RADIOLOGY TRAINING

The use of EMT to benefit radiology training and diagnosis is not postulated as a new idea. Significant research has been conducted to evaluate EMT technologies and systems, to compare and contrast novice and expert readers, and to evaluate metric performance for various anatomical studies and image modalities. The authors do not present a comprehensive survey of all prior research in this area. However, we highlight relevant prior research in the literature that demonstrates the utility and benefits of EMT for radiology training.

Radiology training in the United States is an arduous process which typically involves four years in medical school followed by a one-year internship, four years of residency, and one to three years of fellowship. There are approximately 4500 Diagnostic and Interventional Radiology residents in the United States, all of whom are developing their skills to achieve expertise in imaging diagnosis. Eye motions in interpretation of imaging studies consist of maintained gazes during which visual information is processed (fixations) and rapid gaze shifts between fixations where there is no visual processing (saccades). Together the sequencing of fixations and saccades as a reader interprets a study is known as a "scanpath" as coined by Noton et al. (1971). EMT essentially provides a recording of these scanpaths. In this paper, we use the terms scanpath and search patterns interchangeably. Explorations into EMT within the realm of radiology have been primarily focused on assessment and review. Bertram et al. (2016) found EMT could be used to reflect expertise in interpretation of abdominal CT exams with salient EM differences existing between experienced and non-experienced readers with regard to fixation and saccade durations, as well as reporting adaptive EM reactions to findings as indication of expertise. In addition, there have been several studies looking into the search patterns of experts using eye motion tracking particularly as it pertains to CT and mammography; however, there is a lack of research aimed at teaching radiology trainees to achieve expertise quicker with EMT.

Prior examinations of EMT Technology for Radiology

EMT and Radiology Search Pattern Characterization

Multiple studies have been performed that can characterize search patterns. These studies have shown that there are differences between a novice search pattern and an expert search pattern. The main difference between the two characterizations is the amount of time fixated on a region of interest and the number of fixations. These studies showed that novice readers tend to spend more time fixated on an object and have more fixations per minute than an expert reader.

Kelehan et al. (2019) asserted that a radiologist's search pattern can directly influence patient management, as missed findings can lead to a diagnostic error or missed opportunity for intervention. In this study, search patterns were described to support future studies related to the effects of fatigue, interruptions, and computer-aided detection

of the search patterns and resultant radiology findings. This study quantified the search patterns of multiple radiologists evaluating abdomen and pelvis CT exams. This study found that the inter-reader search patterns between the individuals varied, but the intra-reader search pattern of each individual stayed consistent over the set of scans. The results showed that attendings had fewer number of fixations per second than trainees, which suggested the attendings were more efficient and/or more confident in their search patterns because of their expertise. As part of the study, individuals were categorized as either drillers or scanners based on their search pattern. Qualitative and quantitative metrics and results showed that most radiologists are categorized as drillers. This study provided evidence of discrepancies between the search patterns of expert radiologists and novice readers.

Lévêque et al. (2020) performed a large study to characterize how radiologists perform search, detection, and recognition in screening mammography. In this study, seven expert radiologists representing a range of two to twenty-five years of experience in mammography interpretation evaluated a large number of images, in order to achieve a statistically significant result. This study used both traditional metrics and saccade amplitude in the evaluation of the results and showed that junior radiologists were fixated on the images longer than the senior radiologists. In addition, the results showed that radiologists focus their gaze on small regions of interest when reading scans.

The Cognitive Aspect of EMT Technologies

Brunye et al. (2019) discuss the process for medical interpretation and diagnosis – an interplay between visual perception and multiple cognitive processes. While reading and interpreting medical scans, radiologists must perform visual search and pattern matching, hypothesis generation and testing, and reasoning and problem solving. The authors assert that many expert radiology readers exhibit similar search patterns. As trainees progress through their training regimen and develop proficiency interpreting images, their individual search patterns increasingly resemble expert search patterns. The authors suggest that allowing students to view expert reader's search patterns could facilitate faster progression and learning, similar to novice gamers learning from expert gamers. This is called Eye Movement Modeling Examples (EMMEs) and can include eye gaze information accompanied with audio narrative. A finding in neuroscience suggests these EMMEs can aid students in incorporating these viewed search patterns and behaviors into their practice based on the brain's "mirror system" or mimicry of the observed behavior.

Holistic Assessment

Wu and Wolfe (2019) performed a thorough review of eye tracking and eye movement in the medical domain. The authors assert that viewing eye movements of the reader can identify weak spots that need to be addressed. These results show aspects of the reader's search pattern that need to be improved. Eye tracking could also be used to evaluate the reader's search pattern to evaluate if the reader is progressing and implementing an improved search pattern. In addition, new technologies are emerging and need to be evaluated. These include the inclusion of color in medical imagery and the 3D volumetric imaging. Best search practices for these new visual techniques will need to be created and therefore studies need to be performed to characterize how readers evaluate these studies. The authors state "as long as the advances in medical imaging create new stimuli to improve medical interpretation, there are always new scientific questions that need to be investigated" (Wu & Wolfe, 2019, p.11). With each new medical imaging breakthrough and improvement, training will need to occur to properly use and account for these improvements. Reducing training time while still maintaining education excellence is an important aspect that will need to be addressed. Passive and active use of EMT technology and solutions has the potential to provide performance improvements to the radiology training industry.

Our Solution – An EMT-Based Radiology Training System

Previous research has been performed on the different characteristics between a novice and expert search pattern and the effects of allowing trainees to view expert eye gaze information and search patterns to promote mimicry and faster training progression. Our proposed solution focuses on providing real time evaluation and feedback of a trainee's search pattern by an attending expert radiologist. It is our assertion that this type of training and evaluation can be used to promote confidence in the trainee reading the medical imagery and allow the trainer to better understand the trainee's current skill level to better adapt future training sessions. This type of solution also allows for a more personalized approach to training and evaluation of radiology readers, such as allowing the attending

radiologist to point out mistakes or improvements to the trainee's search pattern in real time to provide personalized training.

Our solution is an integrated radiology training system that uses EMT and expert attendant verbal queueing to decrease the time required to increase competency for the trainees and improved understanding of trainee learning progression for the attending radiologist. During Phases 1 and 2 of our research, we created and tested an initial prototype system that provided promising, but qualitative results. Phase 3 focused on a small study of the prototype that included radiology students and teachers as evaluators of the students. This study included a questionnaire for both the trainees and the attending trainers. This paper highlights tools, methods, and results of Phases 1, 2, and 3.

Phase 1: Creation of an Integrated Prototype

The first step of the creation of the integrated EMT platform prototype was the evaluation of the user needs of the system. Upon the evaluation of the user needs, it was determined that the prototype would need to track the reader's eye gaze, record the eye gaze, and output both a real time view of the user's gaze and a video that could be replayed and examined after the study.

Multiple eye tracking solutions were evaluated, but ultimately the Tobii Eye Tracker 5 was selected because of its out-of-the-box capabilities and low price point when compared to other EMT offerings. Open Broadcaster Software (OBS) Studio 27.1.3 was also selected to record and save video of the EMT study.

In the final integrated prototype, the Tobii Eye Tracker 5 was horizontally mounted to the lower edge of a 24-inch monitor. The computer used was a Dell Precision 7535 and this computer performed all processing of the software locally. The prototype also used Tobii Ghost 1.10.1 and Tobii Experience 1.81 to visualize the output of the Tobii Eye Tracker 5 technology. OBS was also included to record and save the output of the EMT.

Phase 2: Evaluating the Integrated Prototype

The prototype was formatively evaluated as part of a human factors engineering and usability effort. As part of this evaluation, qualitative evaluations of usability objectives were captured and evaluated. This evaluation provided useful information on what aspects of the system should be reworked to provide a more useful solution. This evaluation focused on both the reader and the trainer perspectives and assessment of the system.

First, medical images were selected to evaluate the integrated prototype. Ten medical images were selected by a board-certified radiologist. The selected medical images included bone radiographs, digital subtraction angiograms, and axial CT images. Each of these images were hand-selected based on their outstanding image quality and exported to PDFs that contained no patient information. This set of images also spanned multiple modalities and demonstrated findings that ranged from obvious to subtle, in order to provide an ideal sampling for novice readers to test the EMT technology. For the evaluation, the PDFs of the medical images were viewed using a PDF viewer.

The first participants – two board-certified radiologists and two non-medical, novice readers tested the system. For the evaluation, two individuals would participate at one time. One individual would play the role of the reader and one would play the role of the trainer. For this study, the two board-certified radiologists played the role of the trainer. The role of the reader was performed by either one of the novice readers who had no experience in reading medical imagery or one of the board-certified radiologists. The reader sat in front of the EMT-enabled monitor and read the scan while the trainer evaluated the reader's eye gaze on the viewer platform laptop and provided feedback and verbal coaching in real time. Figure 1, below, shows a novice reader reading the scan while an expert, board-certified radiologist is making corrections to the reader's search pattern and describing the imagery in real time.
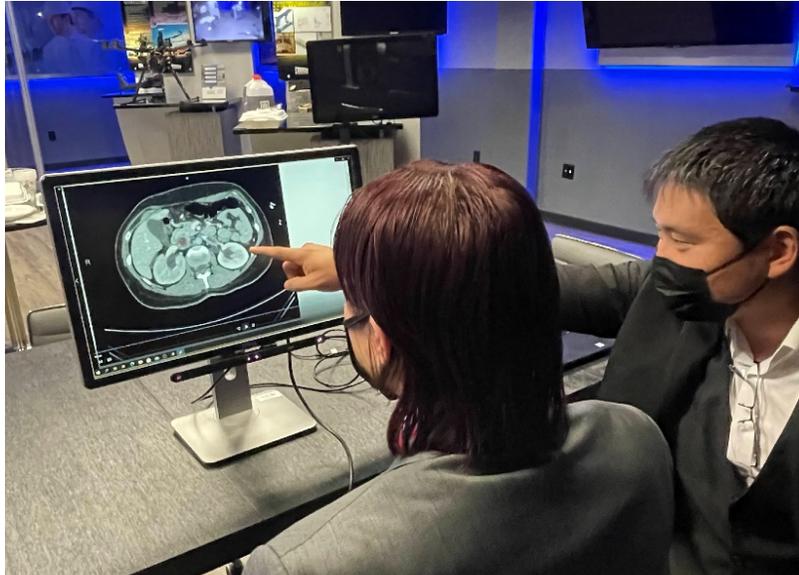
Figure 1 Photograph of the EMT desktop platform with novice reader (Matthew Schultz) whose gaze is fixated at a clot filled inferior vena cava (red bubble on the monitor) and Radiologist (Dr. Junjian Huang) describing what Mr. Schultz is seeing to him.

Figure 2 shows the output of the reader's gaze on a medical image. The saliency map shows the intensity of the gaze. The red on the image indicates that the viewer looked at that part of the image for a longer amount of time than the blue or green indications.
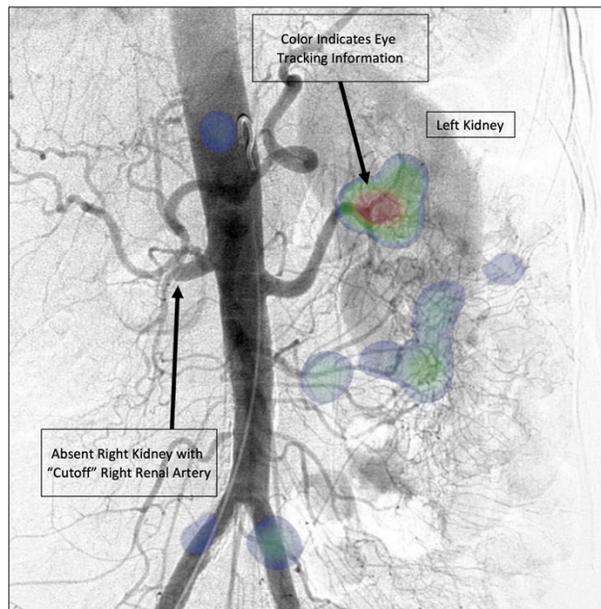

Figure 2 Screen capture of a saliency map using the EMT platform. The image is that of a post traumatic transection of the right renal artery and the gaze of the viewer is currently "fixed" on the left kidney as demarcated by the red heat map.

Initial results indicate that temporal and spatial resolution of the eye gaze tracking system were qualitatively satisfactory when the system was pre-calibrated to the reader. Spatial resolution was degraded when a non-calibrated user attempted to use the system. Using verbal cues, expert attending radiologists were able to guide novice readers to correctly recognize the intended anatomical target. The attending radiologists used the viewer laptop to verify the novice readers were gazing at the intended targets.

Phase 3: Formative Evaluation

After the integrated prototype was formatively evaluated and tested for feasibility, the next phase was to perform a more thorough formative evaluation of EMT technology in relation to radiology training.

The EMT platform evaluation in Phase 2 had several study limitations and needed to be reevaluated for the next formative evaluation. One limitation was the use of a PDF reader to view the images instead of a traditional DICOM viewer. DICOM viewers typically allow the reader to scroll through the images and are optimized for the display of DICOM information, which is the universal format of medical images in radiology. While the feasibility study in Phase 2 was not affected by this limitation as only one image was evaluated from a DICOM study, the formative evaluation performed in Phase 3 would require the reading of multiple sequential images in each imaging exam. Therefore, a Windows based DICOM viewer, RadiAnt, was selected to display the DICOM files.

Another limitation stemmed from the use of the Tobii 5 system which, while affordable and ubiquitous within the gaming industry, was unsuitable for medical related activity. Thus, a new eye tracker was selected, the Gazepoint GP3. Unlike the Tobii 5, the Gazepoint GP3 eye tracker can stand by itself and does not need to be mounted to a monitor. For the formative evaluation, the Gazepoint GP3 was positioned directly underneath the monitor on the table. In addition, the Tobii Ghost software used previously was replaced with the Gazepoint Analysis software that provided recording, gaze overlay visualization, and playback capabilities. The same Dell Precision 7535 computer was used, and this computer performed all processing of the software locally. The prototype platform designed in Phase 2 did not allow for voice recordings. For the formative evaluation in Phase 3, a microphone was positioned in front of the trainer and trainee in order to allow for voice recording during the reading of the medical imagery. Audacity® software was used to record the audio from the microphone.

A lead board-certified academic radiologist from the University of Alabama at Birmingham (UAB) selected the physicians to perform the formative evaluation of the EMT prototype platform. Three groups were identified for this evaluation: (1) a group of academic radiology faculty who would perform the role of attendings or trainers, (2) a group of radiology residents serving as trainees, and (3) a group of non-radiology trainees. The UAB study participants had no serious vision, hearing, or speech pathologies rendering them unable to participate in EMT. The radiology trainees had variable years of experience. The group of non-radiology trainees included personnel with no formal training in a radiology residency program.

One medical imaging study was selected for the formative evaluation by each trainer. This medical images were anonymized and included a singular pathology. For the formative evaluation, the trainee was allowed to scroll through the images within the study using the RadiAnt DICOM viewer. However, image windowing and other image manipulation controls were not allowed during this evaluation. The medical images were selected so that the diagnosis could be reached without these controls. Each of the trainers identified for the formative evaluation had subspecialty training for the selecting medical imaging exam.

At the beginning of every EMT session, the trainee and trainer were given a quick overview of the process and the Gazepoint GP3 eye tracker was calibrated based on the position of the trainee and his or her eyes. After the calibration of eye tracker, the screen and voice recordings were started. The trainee was then allowed to evaluate the medical images by scrolling through the image. While the trainee was evaluating the scans, the trainer was watching the trainee's eye gaze in real-time. In some instances, or as needed, the trainer provided real-time comments or direction to the trainee. There was no time limit for the evaluation; each trainee was allowed as much time as they needed to make a formal diagnosis of the DICOM study. Once the diagnosis was made and discussed with the trainer, the trainee was asked to look or trace specific anatomic structures found in the medical imagery. These directions included fixating on the appendix, scanning the liver, and tracing the bowel in the image. The trainer would again evaluate the trainee based on the trainee's eye gaze information overlayed on the medical imagery. Figure 3 below shows the trainee's gaze as they were tracing the colon in the image. The trainer would provide real-time direction as necessary. Once all of the traces were completed, the eye tracking video and the audio recording were stopped. Next, the video of the trainee's eye gaze overlay video was played back so the trainee could see how they evaluated the imaging exam. The trainer provided additional feedback and suggestions for future improvement while the video played. After the playback and discussions were concluded, the trainee completed a

ten question survey. The trainer also completed a twenty question survey that was based on their overall experience with each of the trainee sessions he or she monitored.
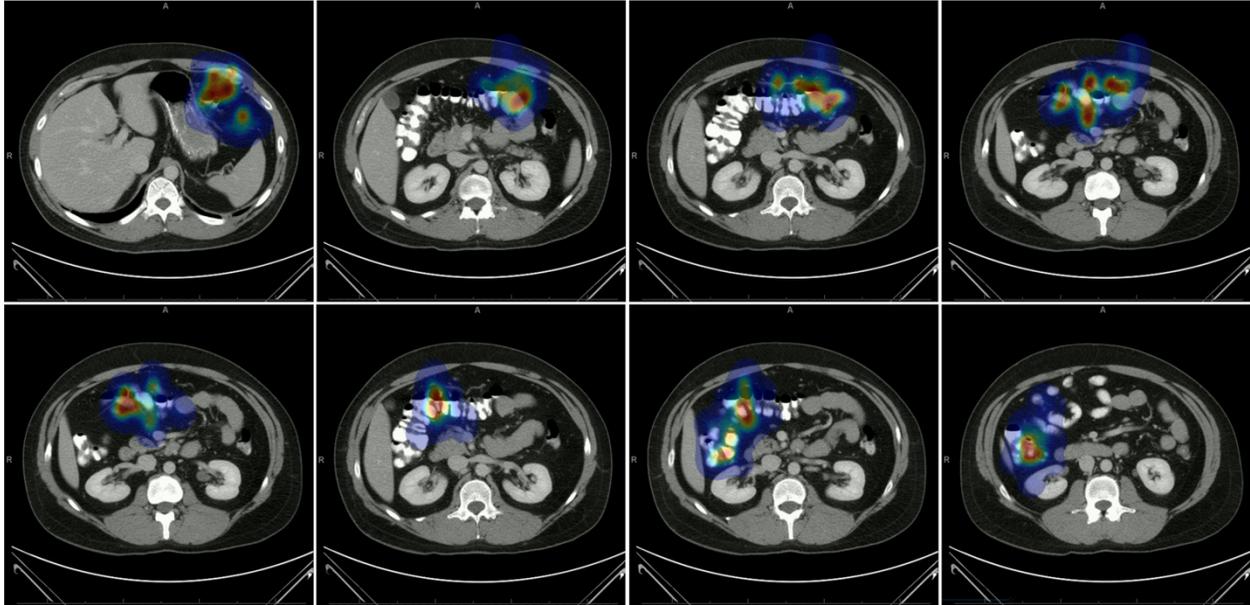

Figure 3 Screen capture of multiple abdominal CT scan images from the DICOM study used in the formative evaluation. The gaze of the viewer is currently tracing the colon as demarcated by the red heat map on each of the images.

Results

Since this was a pilot study requiting an iterative evaluation of an evolving EMT platform, no formal power calculations were performed to establish a sample size. A total of nine physicians participated in this study, including three attending radiologists and six trainees. Trainees enrolled in this study were between their first and fourth year of radiology training: two first years, two second years, and two fourth years.

The study was performed during a single session where trainees underwent eye tracking calibration prior to free-form exam review followed by focused tasks directed by the attending radiologists. Formative evaluation surveys were completed following each session by the trainee/resident whereas attending radiologists completed their surveys following cumulative review of the system.

Trainee Feedback

At the conclusion of every session, the trainee was asked to fill out a short survey which included multiple-choice and free text questions. The multiple-choice questions used a 5-point Likert rating scale ranging from strongly agree to strongly disagree. The figure below shows the results of each of the multiple-choice questions (Figure 4). In all, residents were very intrigued by the EMT solution and how it could be applied to assist in radiology training. Residents reported no negative effect of the EMT system on following attending instructions during the session. All junior trainees agreed that EMT could improve their development or modification of search patterns. In addition, 5 out of 6 residents agreed that EMT could accelerate development of confidence in their reading skills. There was a single trainee who indicated that the eye tracking field for the tracker impeded their ability to read the images. All agreed that EMT would accelerate the attending's confidence in the trainee's reading skills.

Interestingly, the more senior residents indicated that the use of EMT would not accelerate their reading capabilities. This fact coupled with the observation that the senior residents took significantly less time to read the scans and already had an established search pattern indicates that this type of technology is more beneficial to junior residents who are still defining their search patterns. This determination was reiterated by the free response questions included in the survey. One of these questions asked the residents which group would benefit the most from the incorporation

of the EMT system, 67% of the trainees identified junior residents as the primary beneficiaries for EMT technology with the belief it could help benefit search pattern development. The remaining residents cited quality control and trainers as the main beneficiaries of the technology with both citing confidence and accuracy as the main benefit to these groups. Two improvements to the system were also desired by the residents. These included the inclusion of an artificial intelligence (AI) algorithm for image finding detection to show the reader their blind spots and having the system set as a standard of care.
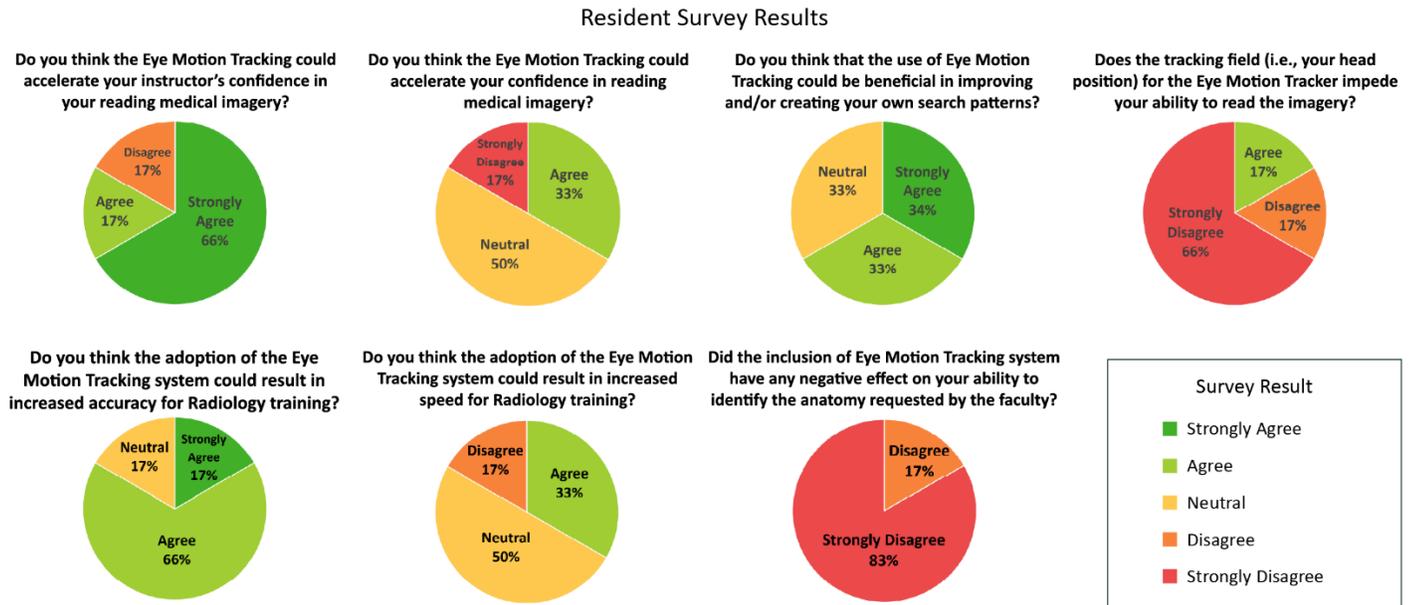

Figure 4 Resident Survey Results

Faculty Attending Feedback

For the study, attendings evaluated at least one resident with some attendings evaluating multiple residents. However, the attending only filled out one survey at the conclusion of the evaluation of the EMT solution. This survey included both multiple choice and free text questions. The multiple-choice questions used a 5-point Likert rating scale ranging from strongly agree to strongly disagree. The figure below shows a synopsis of the results based on EMT characteristic and ratings (Figure 5). Faculty (N=3) were in full agreement that the EMT system was easy to set up, calibrate, and use with no negative impact on trainee's ability to follow instructions. All attendings agreed that the output of the EMT solution is easily understandable and would make the process of training residents easier. In addition, the attendings stated that the inclusion of EMT technologies in radiology training could potentially result in increased efficiency of reads and increased confidence in the attending's ability to gauge the trainee's reading ability. In addition, the attendings agreed that the inclusion of EMT in radiology training could improve and aid in the development of search patterns with each attending seeing the benefit of EMT for training purposes. Two of three attendings noted that the EMT field posed a potential issue as well as the presence of eyewear such as glasses. Once the trainee lost EMT tracking, the attending could no longer view the trainee's eye gaze information. For the specific brand of EMT tracker used during our Phase 3 study, the tracking field was very narrow. This was a concern for most of the attending faculty. Research and technology advances will likely be made to alleviate this issue in the future. All attendings were able to observe trainees complete the specific tasks of fixating on the appendix and terminal ileum, trace the bowel and right ureter, as well as scan the liver. All attendings were able to differentiate between fixations and saccades of the trainee's gaze with this EMT setup. Subjectively, attendings identified trainees, particularly R1's as the primary group that would benefit from this tech, which closely resembles the feedback from the residents
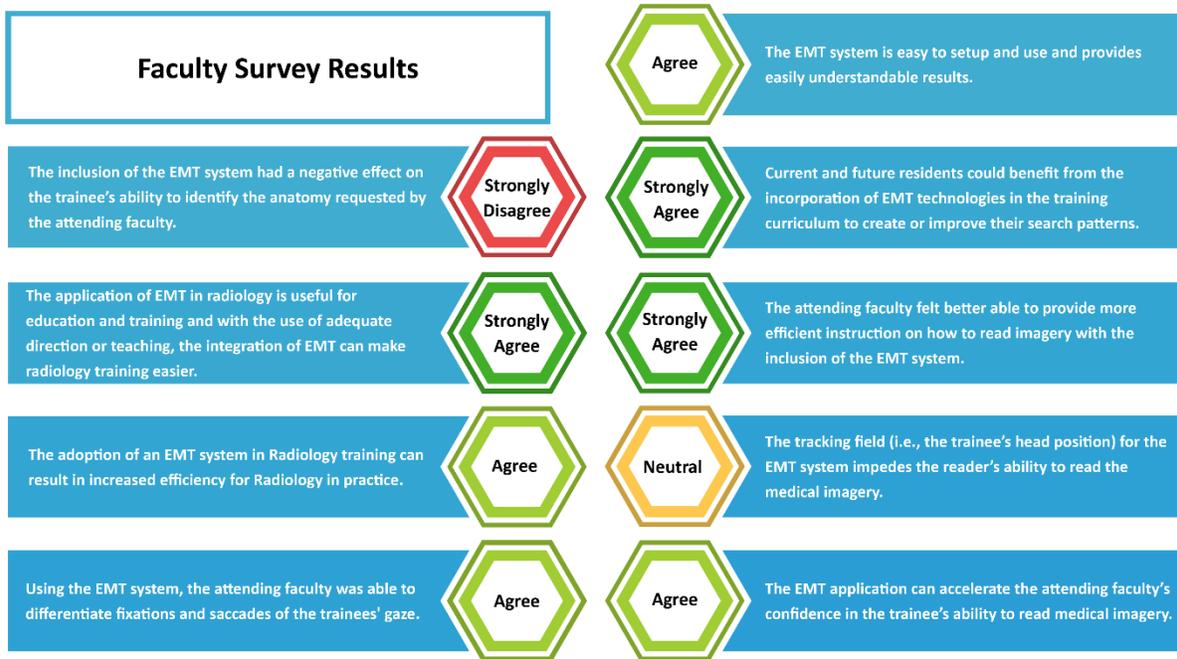
Faculty Survey Results

Agree — The EMT system is easy to setup and use and provides easily understandable results.

The inclusion of the EMT system had a negative effect on the trainee's ability to identify the anatomy requested by the attending faculty. — Strongly Disagree

Strongly Agree — Current and future residents could benefit from the incorporation of EMT technologies in the training curriculum to create or improve their search patterns.

The application of EMT in radiology is useful for education and training and with the use of adequate direction or teaching, the integration of EMT can make radiology training easier. — Strongly Agree

Strongly Agree — The attending faculty felt better able to provide more efficient instruction on how to read imagery with the inclusion of the EMT system.

The adoption of an EMT system in Radiology training can result in increased efficiency for Radiology in practice. — Agree

Neutral — The tracking field (i.e., the trainee's head position) for the EMT system impedes the reader's ability to read the medical imagery.

Using the EMT system, the attending faculty was able to differentiate fixations and saccades of the trainees' gaze. — Agree

Agree — The EMT application can accelerate the attending faculty's confidence in the trainee's ability to read medical imagery.

Figure 5 Synopsis of the Faculty/Attending Survey results showing the average answer for each of the topic areas.

Qualitative Observations

Although this formative evaluation was not designed to assess this, the attendings and observers (engineers present) noted that the way the trainees approached the study was ostensibly different. Junior trainees had a very rigorous and thorough scan path while senior level residents identified the positive finding very quickly, sometimes within seconds of starting their read, compared to their junior colleagues. Also, the scan paths these junior trainees utilize closely mimic the standardized reports that are used within the practice.

There were two non-radiology personnel present during this evaluation – a software engineer who was watching the tracings and listening to the instructions and a registered nurse who was listening to the instructions but not watching the tracings. Both were asked to look at the study and follow the radiologist's instructions. The engineer was able to closely replicate what he was observing the trainees perform with eye tracking, but the nurse was not able to replicate based off of verbal cuing only. This result is particularly interesting, as it mimics the assertions of the brain's mirroring system aspect which has been the subject of prior research (Brunye et al., 2019). Upon evaluation of the EMMEs of the trained residents, the non-radiology-trained reader was able to reliably replicate the results of the trained residents. In contrast, the individual who did not have access to the EMMEs was unable to replicate the results of the trained residents.

Conclusions and Future Directions

This formative evaluation of the EMT-enhanced system demonstrates that the current system assembled by the authors is very close to being fully functional with work needed on improving the eye tracking field. The initial responses from faculty and residents are overwhelmingly positive regarding the potential use of EMT to improve radiology education, particularly amongst junior-level trainees. This coincides well with the purpose of developing this technology within the space of radiology education as the primary goal of this is to accelerate the learning process of novice readers.

The results of this formative evaluation combined with the prior research show that this solution has a high potential to increase trainee and trainer confidence in the trainee's reading ability. This system has also proven through informal means that it can be used to promote the brain's mirroring system to allow for a quicker evaluation and

development of efficient search patterns. These qualitative results are very promising. More thorough and formal evaluations will need to be performed in the future to provide more quantitative results.

The next phase of this project is to develop a call readiness examination with EMT enhancements to gauge progress as well as collect feedback eventually leading to a prospective randomized trial comparing EMT-enhanced education to traditional radiology teaching.

CONCLUSIONS AND FUTURE WORK

In addition to the challenges associated with technology adoption for training, the United States continues to face a growing shortage of physicians. This long-term trend, coupled with the pandemic and increasing image demands, have led to a highly competitive radiologist employment market at a time when medical practices find themselves short-staffed, overworked, and in need of relief. Therefore, training new radiologists quickly and efficiently is of great importance.

In this paper, the authors outlined an integrated EMT platform designed to reduce the training time needed to transition radiology residents from novice to expert readers. Our formative evaluation qualitatively showed that an EMT platform for viewing real-time radiology trainer's eye gaze when reading scans has the potential to benefit radiology training. Approximately 78% of the residents and faculty surveyed cited trainees as the group who could benefit the most from this type of technology. The results also showed that physicians propose this type of technology could improve quality control measures and build confidence in a trainee's ability as a radiology reader. A reduction in time for residents to achieve scan reading competency can begin to mitigate the increasing demands on the imaging community.

There are still areas of improvement to be made and research to be done. The use of AI has become more mainstream over the last decade and could become a game changer in the medical community. The inclusion of an AI agent with an EMT platform could help train students by showing the trainee their blind spots as well as more efficient search patterns. An AI agent could also be used as a second reader to alleviate radiology fatigue and identify a potential misdiagnosis. There are many areas to research including the EMT technology itself, effective User Interface design, and cognitive psychology and how it affects EMT technology solutions. As we expand the system requirements and capabilities, we'll continue to leverage associated research to provide the best integrated solution.

REFERENCES

Bertram, R., Kaakinen, J., Bensch,F., Helle, L., Lantto, E., Niemi, P., & Lundbom, N. (2016). "Eye Movements of Radiologists Reflect Expertise in CT Study Interpretation: A Potential Tool to Measure Resident Development." *Radiology*.

Blanken, L., Lepore, J., & Rodriguez, S. (2018), "America's Military is Choking on Old Technology." *Foreign Policy*.

Brunyé, T., Drew, T., Weaver, D., & Elmore, J. (2019). "A review of eye tracking for understanding and improving diagnostic interpretation." *Cognitive Research: Principles and Implications*.

Fu, H., Wei, Y., Camastra, F., Arico, P., & Sheng, H. (2016). "Advances in Eye Tracking Technology: Theory, Algorithms, and Applications." *Computational Intelligence and Neuroscience*.

Kelehan, L., Fong, A., Blumenthal, J., Kandaswamy, S., Ratwani, R., & Filice, R. (2019). "The Radiologist's Gaze: Mapping Three-Dimensional Visual Search in Computed Tomography of the Abdomen and Pelvis." *Journal of Digital Imaging*.

Lévêque, L., Yong, P., & Liu, H. (2020). "Studying the Gaze Patterns of Expert Radiologists in Screening Mammography: A Case Study with Breast Test Wales." 28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands.

Noton D. & Stark L. (1971). "Scanpaths in eye movements during pattern perception." *Science*. doi: 10.1126/science.171.3968.308. PMID: 5538847

Valliappan, N., Dai, N., Steinberg, E., He, J., Rogers, K., Ramachandran, V., Xu, P., Shojaeizadeh, M., Guo, L., Kohlhoff, K., & Navalpakkam, V. (2020). "Accelerating Eye Movement Research Via Accurate and Affordable Smartphone Eye Tracking." *Nature Communications*.

Visweswaran, S., King, A., Tajgardoon, M., Calzoni, L., Clermont, G., Hochheiser, H., & Cooper, G. (2021). "Evaluation of eye tracking for a decision support application." *JAMIA Open*, Volume 4, Issue 3.

Walter, M. (2021). "ED imaging demand skyrocketed in 2021, and many radiology providers are struggling to keep up. What's next for the specialty?" *Medical Imaging Review*.

Wu, C. & Wolfe, J. (2019), "Eye Movements in Medical Image Perception: A Selective Review of Past, Present and Future." *Vision Special Issue – Eye Movements and Visual Cognition*.