

# Eye Scan-Path Training with Voice Dialog on Chest X-Rays and Analyzing Results using Gridinger's Classification

**Takumi Bolte**  
Clemson University  
100 McAdams Hall  
Clemson, SC 29634  
tbolte@clemson.edu

**Alison Nolan**  
Clemson University  
100 McAdams Hall  
Clemson, SC 29634  
alisonn@g.clemson.edu

**Marvin Andujar**  
Clemson University  
100 McAdams Hall  
Clemson, SC 29634  
manduja@g.clemson.edu

## ABSTRACT

This paper describes the applicability of a novel classification method to discriminate eye-tracking data of chest x-ray images. Using this on training data, significant differences in scanpaths can be determined between pre and post training. It is concluded that providing training can change the viewing patterns of a subject and how they examine x-ray imaging.

## Author Keywords

Eye-tracking training techniques; Grindinger's classification; eye scan-path training; chest x-rays

## ACM Classification Keywords

H.1 Models and Principles: User/Machine Systems; H.5 Information Interfaces and Presentation: User Interfaces

## General Terms

Human Factors

## INTRODUCTION

The number of x-ray procedures each year is increasing at a rate of about 5.5% [9]. There were an estimated 182.9 million x-rays taken in the year 2010. Up to 20% of these images are incorrectly read, meaning there was a "miss" or a false positive. A miss implies that there is an abnormality present, but the radiologist fails to notice it. A false positive describes the situation where a radiologist classifies a normal image as abnormal [10]. Errors are clearly a problem, and this paper describes a method to help train radiologists to read images more accurately.

Eye-tracking is used to help train novice doctors. There are many studies that have implemented eye scan-path training techniques. The typical experiment recruits novice doctors to read x-ray images. The novices are first asked to read the image without any additional support. They are then shown an image with the eye scan-path of an expert radiologist. After exposure, it is found that the novice radiologists can better identify pulmonary nodules. Despite the increase in

performance, there is still a certain level of error from the novices [6]. This paper examines the addition of a voice-over to improve the accuracy from the novices.

It is expected that the novice doctors will improve both the speed at which they detect anomalies, and also the precision if they have correctly identified it. The addition of speech to the training should help the novices better understand what the experts are searching for in the chest x-ray. The novices were measured in their speed, correctness of the diagnosis, and eye scan-path. The speed and correctness were simple quantitative data analysis. The eye scan-path analysis was analyzed using Gridinger's classification [1]. The classification compares the differences of the novice's eye scan-path versus an expert's eye scan-path. In addition, novices were asked their opinion on the training method. The qualitative responses are more difficult to analyze, but gives insight to a user's opinion of the training.

## RELATED WORK

There are several studies conducted in eye tracking on chest x-rays. Sadasivan et. al. used eye tracking to record the point of regard of an expert inspector while performing an inspection task in a virtual reality simulator. Their results suggests that trained novices should adopt a slower paced strategy through increased fixation duration, which they learn a more thoughtful target search strategy [2]. Vitak et. al. showed through a study with a scripted/recorded video of an expert's think-aloud session, which was augmented by an animation of their scan-paths, can result in an effective aid for learners of visual search [3]. A similar concept was Dempere-Marco et. al. work. They presented a conceptual framework where they gathered the expert knowledge in lung radiology to be implemented in a practical application [4]. This framework may be useful to be applied for chest x-rays training. Dempere-Marco et. al. had another work where the spatial characteristics of visual search were studied. They combined the intrinsic visual features of the fixation points for comparing different visual search strategies. Their study showed through both spatial and feature space representation, it is possible to untangle the uncorrelated fixation distribution patterns to reveal common visual search behaviors [5]. Litchfield et. al. reported a study whether an experienced and inexperienced radiographer would benefit from knowing where another

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*Wherever this will be submitted information*

person looked during pulmonary nodule detection. Their results showed improvement is maintained whether radiographers are shown the eye movement patterns of an expert or a novice [6].

There is also work dedicated to the analysis of eye tracking data. Beeks et. al. showed the analytical computation of the Gaussian Quadratic Form Distance and evaluated its retrieval performance by making use of different benchmark image databases [7]. Coen worked on a new metric to measure similarity between spatial probability distributions. In his work, he tested it in multiple ways to demonstrate its efficiency [8].

Despite the amount of work in this area, previous research has not concentrated on the voice training prior to the task and used the Gridinger's classification. Therefore, this paper presents a study conducted where the participants hear the eye scan-path from an expert prior to performing the task. Also, the Gridinger's classification is used for data analysis.

## METHODOLOGY

This section outlines the specifics of the device used, the procedure followed, information about the subjects, and the experimental design.

### Experimental Design

The experiment was designed as a within subject. All the participants performed the same task. Five images were given first to determine their ability to do scanpath, then the training was played. Lastly, another six images were shown to evaluate their new knowledge after the training.

### Apparatus

The eye movements were captured using the GP3 Desktop Eye-Tracker by GazePoint. This is a relatively new off the shelf eye-tracking device. It has an accuracy of 0.5 to 1 degree of visual angle. Its sampling rate is of 60Hz. The movement for both vertical and horizontal is 25cm x 11cm. This device works with 22-inch displays or smaller.

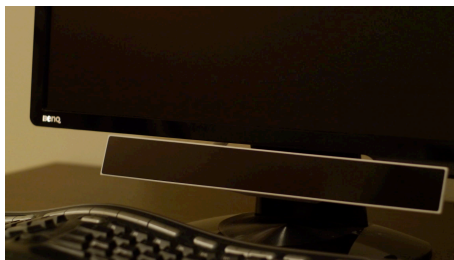


Figure 1. GP3 Desktop Eye-Tracker

### Stimulus

11 stimuli images were selected by an expert doctor (figure 2). The images contained a subset of both healthy and maladed CXRs. The purpose of the images was to be scanned by the novice doctors to identify the problems if

any. The novice doctors were able scan each image for 10 seconds.

### Participants

There were a total of 11 participants (females: 7 males: 4) with ages ranging from 18 – 37. The participants were students from Clemson University. The subjects did not have any experience on reading chest x-rays. The participants were asked about the value of having the training video. About half (5 people) responded that the video was "not helpful", 5 reported "somewhat helpful", and one believed it to be "extremely helpful." Only one participant provided additional feedback. They wrote, "I could barely hear the instructor on the video, nor could I remember the majority of the instructions after the first 5 or so."

### Procedure

The participants filled out a pre-assessment assessing them on their previous knowledge on scanning chest x-rays. Then, they are given five CXR images and asked simply to examine each one. Then, the participants watch a tutorial video explaining the proper way to read chest x-rays. After this training session, the participants are given six chest x-rays. There is not a specific order the participants scan the chest x-ray.

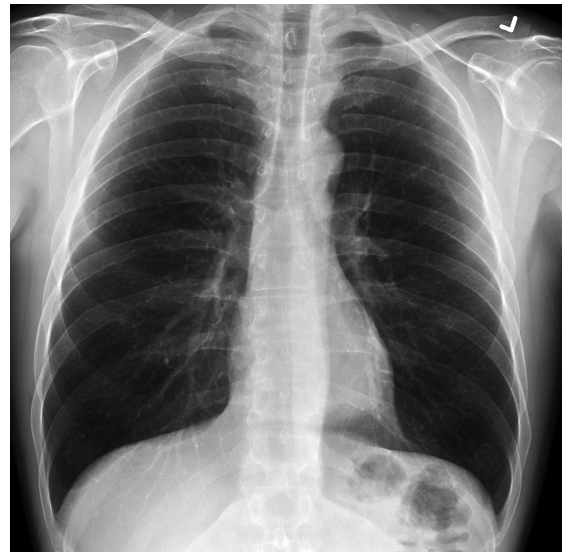


Figure 2. An example of a normal chest x-ray

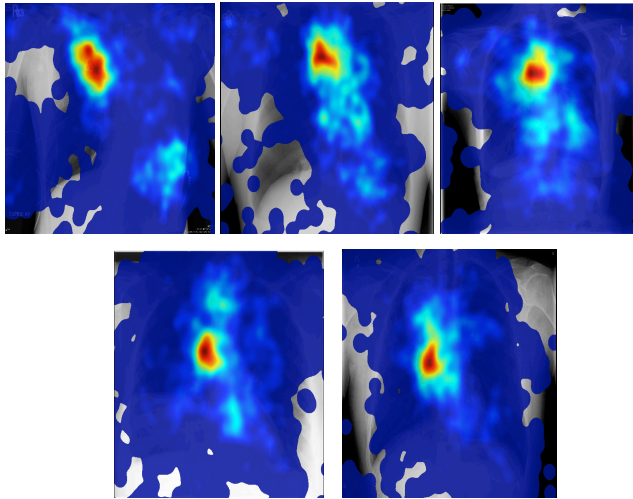
## RESULTS

Each subject had both a pre and post training data set. Using Grindinger's classification, a positive and negative training set is used to determine significance between eye tracking data. The aggregate data for pre and post training were used as the positive and negative training sets respectively. Classification of pre and post training using Grindinger's metric for scanpath classification clearly demonstrates a dichotomy between these two data sets.

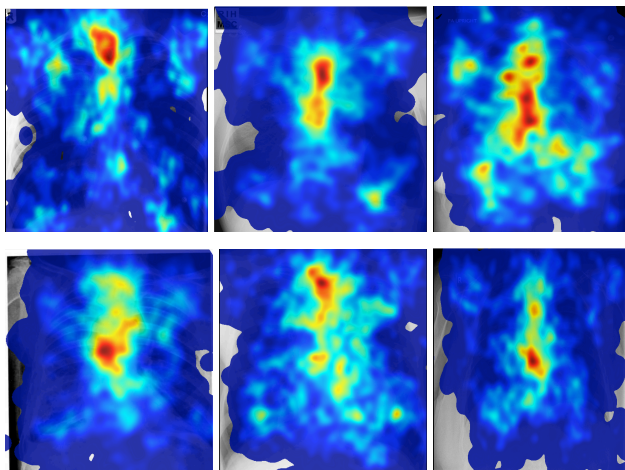
	Pre / Post Training
ACC	1.00
AUC	1.00

**Table 1. ACC and AUC results**

The ACC displays the accuracy that a given set of scanpaths will be classified to a given set, in this case in either pre or post training (Grinding). An ACC value of 1.0 indicates that each set of scanpaths was successfully classified to one of our classifiers, pre training or post training. The AUC value shows a highly discriminable difference between the two training sets. An AUC value > 0.7 suggests high discriminability in regards to classification of a set. With an AUC of 1.0, each set of scanpaths were classified under pre or post training classification. This suggests a significant difference in the way a subject examines an x-ray image after training.



**Figure 3. Aggregate pre training stimuli**



**Figure 4. Aggregate post training stimuli**

Examining both the aggregate heat map visualization for both pre and post training (figure 3 and 4), there is a significant change with the viewing patterns of the subjects. During the pre training phase, the overall heatmap generation displays a distinct fixation in a single focused location on the chest x-ray. Following the training phase, the heatmap shows a shift from a stationary eye position to a spread over the center of the x-ray.

## DISCUSSION

Pre and post training of the subjects demonstrated significant changes of fixations of a chest x-ray, even with novice participants. Grinding's classification metric further benefits this result with providing a high ACC and AUC value, giving both confidence in accurately classifying a pre/post training data set and discriminating these sets.

## SUMMARY/FUTURE WORK

This study was done on a very small scale, having only eleven participants. The next round will include more participants that have experience reading chest x-rays. It will give better insight to see if training helps only novices, or also those who have some experience. The tutorial video was also lengthy. The tutorial can be cut down to a shorter explanation of how to read the x-ray. Participants commented on the amount of information presented as overwhelming.

There are also collaboration opportunities available with the local hospital in Greenville. They can provide more images to have a mix of both abnormal chest x-rays and normal images. The study can be extended to have participants think-aloud; that is, they explain what they are looking at and point out abnormalities. This will give another parameter to measure - accuracy and correctness.

## REFERENCES

1. Grinding, Thomas. *EVENT-DRIVEN SIMILARITY AND CLASSIFICATION OF SCANPATHS*. Diss. Clemson University, 2010. N.p.: n.p., n.d. Print.
2. Sadasivan, S. (2005). Use of eye movements as feedforward training for a synthetic aircraft inspection task. *Proceedings of the SIGCHI conference on Human factors in computing systems – CHI'05* (p. 141-149). New York, New York, USA: ACM Press.
3. Vitak, S. a., Ingram, J. E., Duchowski, A. T., Ellis, S., & Gramopadhye, A. K. (2012). Gaze-augmented think-aloud as an aid to learning. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* (p. 2991-3000). New York, New York, USA: ACM Press.
4. Dempere-Marco, L., Hu, X., & Yang, G.-Z. (2010). A Novel Framework for the Analysis of Eye Movements during Visual Search for Knowledge Gathering. *Cognitive Computation*, 3(1), 206–222.

5. Dempere-Marco, L., Hu, X.-P., Ellis, S. M., Hansell, D. M., & Yang, G.-Z. (2006). Analysis of visual search patterns with EMD metric in normalized anatomical space. *IEEE transactions on medical imaging*, 25(8), 1011–21.
6. Damien Litchfield ; Linden J. Ball ; Tim Donovan ; David J. Manning ; Trevor Crawford; Learning from others: effects of viewing another person's eye movements while searching for chest nodules. Proc. SPIE 6917, Medical Imaging 2008: Image Perception, Observer Performance, and Technology Assessment, 691715 (March 06, 2008).
7. Beecks, C., Ivanescu, A. M., Kirchhoff, S., & Seidl, T. (2011). Modeling image similarity by Gaussian mixture models and the Signature Quadratic Form Distance. In *2011 International Conference on Computer Vision* (pp. 1754–1761). IEEE.
8. Coen, M. (2007). A similarity metric for spatial probability distributions. *submitted to IJCAI*. Retrieved from <http://people.csail.mit.edu/mhcoen/Similarity.pdf>.
9. Prochaska, Gail (2011). IMV reports general x-ray procedures growing. Retrieved from <http://www.prweb.com/releases/2011/2/prweb8127064.htm>.
10. Goddard, P., Leslie, A., Jones, A., Wakeley, C., & Kabala, J. (2001). Error in radiology. *British Radiology*, 74(886), 949-51.