Comparing Visual Search Patterns in Chest X-Ray Diagnostics

Catarina Moreira*
catarina.pintomoreira@data61.csiro.au
Data61, CSIRO
Eveleigh, NSW, Australia

Diogo Alvito
diogoalvito@tecnico.ulisboa.pt
Instituto Superior Técnico / INESC-ID
University of Lisbon
Lisbon, Portugal

Sandra Costa Sousa
Lusíadas Knowledge Center
Lisbon, Portugal
sandra.costa.sousa@lusiadas.pt

Isabel Blanco Nobre
Lusíadas Knowledge Center
Lisbon, Portugal
isabel.blanco.nobre@lusiadas.pt

Chun Ouyang
Queensland University of Technology
Brisbane, Australia
c.ouyang@qut.edu.au

Regis Kopper
University of North Carolina at Greensboro
Greensboro, NC, USA
kopper@uncg.edu

Andrew Duchowski
Clemson University
Clemson, SC, USA
duchowski@clemson.edu

Joaquim Jorge
Instituto Superior Técnico / INESC-ID
University of Lisbon
Lisbon, Portugal
jorgej@acm.org

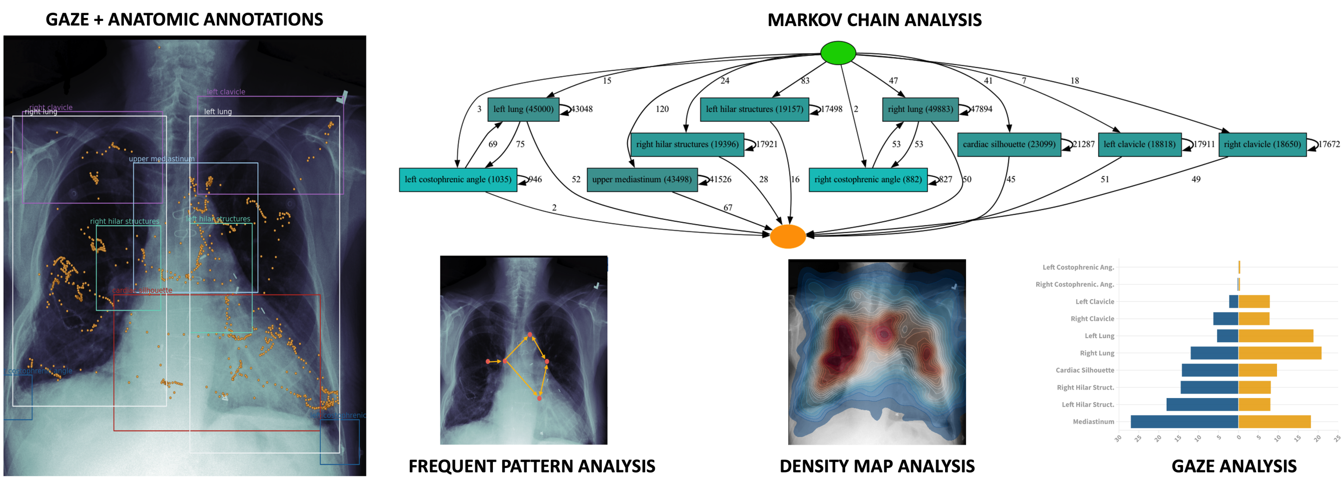

Figure 1: Main contributions of this research: given a dataset containing a radiologist's eye-tracking data and annotations of the anatomical regions of the chest X-ray, (1) we extracted Markov chains corresponding to the radiologist's visiting patterns, (2) we performed a density map analysis to determine the most visited regions in the X-ray, and (3) we conducted a frequent pattern analysis to understand what were the main visitation patterns between different diagnosis.

ABSTRACT

Radiologists are trained professionals who use medical images to obtain clinically relevant information. However, little is known about visual search patterns and strategies radiologists employ during medical image analysis. Thus, there is a current need for guidelines to specify optimal visual search routines commonly used by radiologists. Identifying these features could improve radiologist training and assist radiologists in their work. Our study found that during the moments in which radiologists view chest X-ray images in silence before verbalizing the analysis, they exhibit unique search patterns regardless of the type of disease depicted. Our findings suggest that radiologists' search behaviors can be identified at this stage. However, when radiologists verbally interpret the X-rays, the gaze patterns appear noisy and arbitrary. Current deep-learning approaches train their systems using this noisy and arbitrary gaze

*Also with INESC-ID Lisboa / Instituto Superior Técnico, ULisboa.

ETRA '23, May 30–June 02, 2023, Tubingen, Germany

ACM ISBN 979-8-4007-0150-4/23/05.
https://doi.org/10.1145/3588015.3588403

data. This may explain why previous research still needs to show the superiority of deep-learning models that use eye tracking for disease classification. Our paper investigates these patterns and attempts to uncover the eye-gaze configurations during the different analysis phases.

CCS CONCEPTS

• **Computing methodologies → Tracking**.

KEYWORDS

datasets, neural networks, gaze detection, text tagging

ACM Reference Format:
Catarina Moreira, Diogo Alvito, Sandra Costa Sousa, Isabel Blanco Nobre, Chun Ouyang, Regis Kopper, Andrew Duchowski, and Joaquim Jorge. 2023. Comparing Visual Search Patterns in Chest X-Ray Diagnostics. In *2023 Symposium on Eye Tracking Research and Applications (ETRA '23), May 30–June 02, 2023, Tubingen, Germany*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3588015.3588403

1 INTRODUCTION

Radiologists are in short supply worldwide, with low-income and middle-income countries (LMICs) suffering the most significant shortages [Hricak et al. 2021]. There has been increased interest in finding solutions to mitigate this problem. A potential solution is to automate the diagnostic process using Artificial Intelligence (AI). There are, however, concerns about bias regarding the use of AI in high-stakes decision-making, as well as a lack of transparency in how decisions are made. Using eye-tracking data from radiologists to analyze X-rays is an effective way to solve this problem. Research has been conducted on radiologists' behavior patterns to reduce misclassifications and prepare medical students for the future. Although these methods are promising, the noise in eye-tracking data limits their scalability. Recently, studies have been made to incorporate eye-tracking data and X-ray images in multimodal deep learning architectures to improve prediction accuracy based on the underlying radiologists' behavioral patterns. However, the literature is not conclusive on whether eye-tracking data is helpful for deep learning. Still, most recent work has used simple saccade and fixation patterns. We argue that more elaborate analyses might yield signature patterns to contribute to more precise diagnostics. To this end, we analysed gaze patterns in two distinct moments during a X-ray analysis protocol: (1) when the radiologist is silent at the beginning, and (2) when the radiologist is reporting their findings. We conducted a three-pronged approach focusing, first, on density maps that capture atemporal patterns to reveal the most visited regions of Chest X-Ray (CXR) images. Second, we looked at the temporal visitation patterns using Markov Models to uncover differences over time across different diagnoses. Finally, we adopted string analysis to compare sequences between visited areas. One key finding of our study is that initial (silent) and final (reporting) gaze patterns differ markedly regardless of diagnostic. Our results suggest that information gathered from reporting phase data using these three techniques may well provide the key to usher in more precise machine learning techniques and that indeed Eye Gaze patterns may be leveraged to enhance deep learning approaches.

The remainder of this paper is organized as follows. In the next section, we review relevant literature on the topic. Subsequently, we present the methodology of our study, including the datasets used and a detailed description of the three analyses conducted: Density Maps, Markov Chains, and Longest Repeating Patterns. We conclude with a summary of our findings, and discuss suggestions for future research.

2 RELATED WORK

Interpreting and diagnosing CXR images is a complex task combining visual perception and multiple cognitive processes. Several medical conditions can only be detected through the analysis of medical images, and doctors have to make decisions with images that may not have the best conditions. This highlights the need for approaches to this challenge. Eye Gaze applications have become popular in recent years [Harezlak and Kasprowski 2018; Zammarchi and Conversano 2021].

Understanding the visual search behavior of radiologists when viewing medical images is crucial as their visual search patterns are unique and highly subjective in nature [Alamudun et al. 2015; Carmody et al. 1984; Thomas and Lansdown 1963; Tuddenham and Calvert 1961].

Kundel et al. [Kundel et al. 1990] reported that missed diagnoses in chest radiographs result from various errors, with scanning errors accounting for about 30%, recognition errors accounting for about 25%, and decision-making errors accounting for about 45% of the total errors respectively. The study also identified potential causes of these errors using chest radiographs. These causes include: (i) premature termination of search after the detection of an abnormality, commonly explained as satisfaction in the 'quest for meaning' [Samuel et al. 1995; Tuddenham 1962]; (ii) termination of search when the probability of reporting more false positives increases with increase in search time [Berbaum et al. 1991, 1990]; (iii) premature termination due to lack of confidence in reporting abnormalities55; (iv) limited perceptual resources [Berbaum et al. 1996, 1991]; and (v) visual neglect of certain regions in an image while detecting abnormalities [Berbaum et al. 1996].

Previous studies have demonstrated that the incorporation of virtual reality (VR) technologies in radiology reading rooms offer a practical, adaptable, portable, and financially viable solution to address the limitations of traditional settings. Recent studies by Sousa et al. [Sousa et al. 2017], have shown the potential benefits of VR technologies in enhancing radiologists' performance and improving the diagnostic accuracy. However, using eye-tracking data in evaluation of X-ray images has yet to be fully explored.

Recently, the EYEGAZE dataset [Karargyris et al. 2021] and REFLACX [Bigolin Lanfredi et al. 2022] studies employed eye-tracking technology to track visual search behaviors of radiologists while interpreting medical images. In the EYE GAZE dataset, a radiologist who was American Board of Radiology certified with over five years of experience performed routine radiology readings using the GazePoint GP3 Eye Tracker in multiple sessions, with 30 cases per session over a period of two months. On the other hand, REFLACX involved five board-certified subspecialty-trained thoracic radiologists who closely worked on the study design.

Ganesan [Ganesan et al. 2018] shows that identifying the factors that influence radiologists' visual search behavior can help understand their perception and interpretation of medical images, which in turn could lead to the development and implementation of effective strategies to improve the ability to detect abnormalities. Factors that have been identified as significant influences on radiologists' visual search patterns include expertise, satisfaction of search, visual fatigue, confidence in reporting abnormalities, training received, and prior knowledge. Among these factors, expertise has been identified as having the most significant impact on visual search patterns. Eye-tracking studies have also shown that more experienced readers tend to have more effective visual search strategies, which can be observed through the difference in visual search patterns among readers with different levels of experience.

In medical imaging, radiologists typically establish an overall impression of an image within the first few seconds of viewing, called the global impression, as previously reported [Kundel and Nodine 1975; Kundel et al. 1991; Nodine and Kundel 1987]. After resolving initial perturbations, radiologists carry out a scan of the whole image to detect abnormalities that were not conspicuous enough to be picked out during the initial global impression. This is followed by the confirmation of initial decisions made without focusing on finding new abnormalities. The global-focal search model of visual search proposes the four main components of visual search: (i) Global impression, (ii) Foveal verification, (iii) Discovery scanning, and (iv) Reflective search [Nodine and Kundel 1987].

Some studies have already shown results regarding the improvement of the identification of pathologies based on observing another person's eye movements. For instance, in a study where there were improvements regarding the identification of pulmonary nodules [Litchfield et al. 2010].

3 METHODOLOGY

We analysed gaze patterns in moments where the radiologist is silent and when the radiologist is reporting their findings and conducted a three-pronged approach: (1) focusing on density maps that capture atemporal patterns to reveal the most visited regions of CXR images; (2) investigating temporal visitation patterns using Markov Models to uncover differences in temporal order across different diagnoses; and (3) adopting a string analysis to compare sequences between visited areas.

3.1 Dataset Used

The EYEGAZE [Karargyris et al. 2021] dataset contains eye-tracking information for 1,083 CXR images selected from the MIMIC-CXR dataset [Johnson et al. 2019b,a]. EYEGAZE further contains the final diagnostic of the images as a label, which can be "Normal" (a healthy subject), "Pneumonia", and "Congestive Heart Failure" (CHR). Additionally, EYEGAZE features raw eye gaze x and y coordinates of a single radiologist reading sessions for all images, captured at millisecond intervals. The dataset also contains fixation points due to the post-processing of the raw gaze data. Two of the EYEGAZE dataset's key features are: (1) it provides bounding boxes corresponding to the different anatomical regions of the thorax, and (2) it contains audio transcripts of the radiologist reading sessions. When creating the EYEGAZE dataset, the radiologist was asked to describe a CXR image while their eye gaze patterns were captured and the audio recorded. These data enable us to investigate the differences in the radiologist's search patterns during moments of silence and moments of reporting. These findings can contribute to more refined and human-centered models for diagnosis prediction.

3.2 Density Map Estimation

In order to investigate the correlation between diagnostic outcomes and visual search patterns, we employed a density map analysis of the radiologist's raw gaze data. The density maps were generated by applying Gaussian distributions to cluster the x and y coordinates of gaze points.

We also performed a gaze analysis by computing the percentage of gaze points that fall inside each anatomic region during the silent and reporting diagnostic phases. Note that one gaze point can correspond to several overlapping anatomical regions, e.g., "Right Hilar Structures" can overlap parts of the "Right Lung" region. To address this issue, we applied the smaller-area heuristic, choosing the smallest area anatomical region that contains a given gaze point. The rationale is that the least area should identify the most specific anatomical feature. Other heuristics are possible, such as nearest-neighbors-visited-in-sequence, although it is unclear whether the added complexity would yield significantly better insights.

According to professional radiologists, the process of interpreting medical images involves two phases: the exploration phase and the reporting phase. During the exploration phase, radiologists silently observe the image and identify regions of interest. In the reporting phase, they revisit these regions and make their final diagnostic conclusions. This process aligns with cognitive theories, such as the adaptive gain theory [Aston-Jones and Cohen 2005; Gilzenrat et al. 2010], which posits that the human mind balances the tendencies of exploitation and exploration in visual search. Exploitation refers to focusing on known and high-salience regions that are likely to yield diagnostic information, while exploration refers to the process of scanning the entire image for additional regions of potential diagnostic significance. As radiologists interpret medical images, they engage in both exploitation and exploration, allowing them to make accurate and comprehensive diagnostic conclusions [Brunyé et al. 2016].

3.3 Markov Chain Discovery

Density Maps allow us to identify hot spots in an atemporal manner. However, these blur the sequential visitation patterns, which can provide important insights regarding sequence analysis and possible lesion discovery process. We mapped each visited bounding box into a distinct state in a Markov Chain (MC) to unearth these features. MCs are useful in capturing larger transition patterns across many different images grouped by diagnosis. This enables us to further investigate the exploration/exploitation phases of the adaptive gain theory by analyzing the dynamics of the radiologist's gaze patterns. Again, we present our findings in two phases (silent and reporting moments) as in the previous analysis. We achieved this by simply counting the number of times the gaze of the radiologist transitioned from one annotated anatomical region to another. The transition matrix was then computed by normalizing these values.

We then plotted the Markov Chains and animated how the radiologist's gaze transitions between the different anatomical regions. These animations are provided in the supplementary materials of this work. The animations were produced using a process mining tool from *Fluxicon* called *Disco*.[1] Note that when directly mapping the gaze points to a Markov Chain, this visualization becomes a *spaghetti diagram*, which is very hard to read. In order to provide a visualization that promotes the readability and the understandability of the radiologist's gaze patterns, we plotted the MCs using the 30% most frequent visitation patterns.

3.4 Frequent Pattern Extraction

Finally, we analyzed the most frequent sequential patterns that the radiologist performed for each different diagnosis for the silent and reporting phases. To achieve this, we mapped each anatomical region of the sequence to a unique character, obtaining a string with a sequence of characters that uniquely describe the radiologist's visiting patterns. To find the most frequent patterns in this string, we applied the longest repeating sub-sequence (LRS) algorithm [Hirschberg 1975], used in bio-informatics to find similarities in protein sequences [Gusfield 1997]. We computed the 5 most frequent patterns depicted in Figure 4 as yellow arrows. Note that in each of these Figures, the background CXR images are merely provided for context.

4 RESULTS AND DISCUSSION

In this section, we present the key findings from our study of the EYEGAZE dataset. We provide a detailed analysis of the most noteworthy results and offer insights into the implications of these findings.

4.1 Gaze Density Analysis

The results obtained in the Gaze Density Analysis can be found in Figure 2.

Figure 2 (left) presents the resulting cluster centers of the density maps, indicating the radiologist's most frequently visited regions across all images of the dataset. The intensity of the color on the map ranges from dark blue for the least visited areas to bright red for the most visited spots, with values on a $[0, 1]$ interval.

Our analysis of the radiologists' gaze patterns during moments of silence revealed that the density maps were similar across diagnostic categories. This suggests that the radiologists were engaged in the exploration process, wherein they scanned the X-ray image in search of regions of interest. The density maps further revealed that the radiologists predominantly focused on the upper mediastinum region of the thorax, which encompasses vital structures such as the heart and hilar structures. This observation aligns with the exploration/exploitation model [Aston-Jones and Cohen 2005; Brunyé et al. 2016; Gilzenrat et al. 2010], which posits that radiologists engage in a dynamic balance between exploring the image to identify new information and exploiting previously acquired knowledge to make a diagnosis.

The bar charts in Figure 2 (right) indicate significant variations between the reporting (yellow) and silent (blue) phases across different diagnostic modalities, as evident in both the left and right columns of each row in the data. Furthermore, the bar charts also reveal substantial differences in reporting patterns. During the silence periods, radiologists tend to focus their attention on the mediastinum region. However, during the reporting phase, attention is shifted toward the lung regions, regardless of the specific diagnosis. This suggests that radiologists selectively focus on these regions as they report on their findings. This pattern of attentional allocation is consistent with the exploitation phase of the exploration/exploitation model [Brunyé et al. 2016], in which individuals tend to focus on regions of high relevance or salience. Further research is needed to explore the underlying cognitive and neural mechanisms driving these attentional shifts in radiologists and investigate whether these patterns generalize to other populations and modalities.

4.2 Markov Chains

Figure 3 presents the Markov Chains that were directly extracted from the radiologist's visitation patterns across different diagnoses for the silent and reporting moments. Note that, for visualisation purposes, Figure 3 only presents the 30% most visited paths of the radiologist.

Our overall statistical results indicate that during the Silent phase, the time the radiologist takes to explore the different CXR images is: 2.09sec for Normal images, 2.15sec for Pneumonia images, and 2.11sec for CHF images. This shows that independently of the pathology present in the CXR, the radiologist takes approximately the same time to initially process the image's feature, as predicted in the global impression phase of the four-stage model [Nodine and Kundel 1987]. During the initial phase, the radiologist forms a general impression of the image and quickly identifies the most

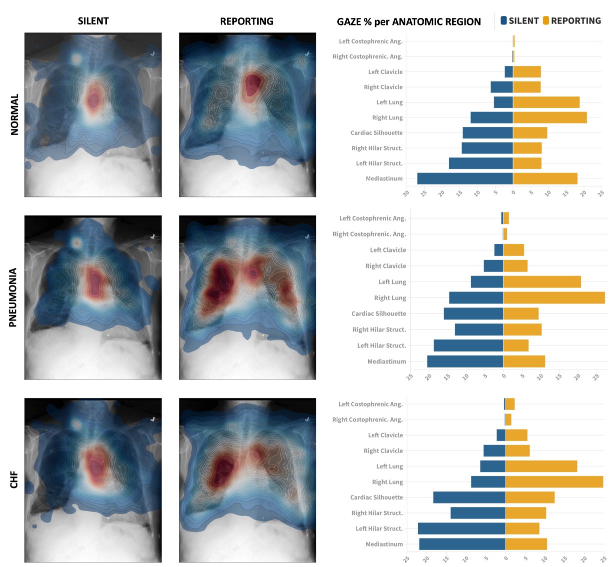

Figure 2: Density maps of the x and y coordinates of the radiologist's gaze across all images of the dataset. Note that the background CXR image is to promote the readability of the density regions.

[1]https://fluxicon.com/disco/

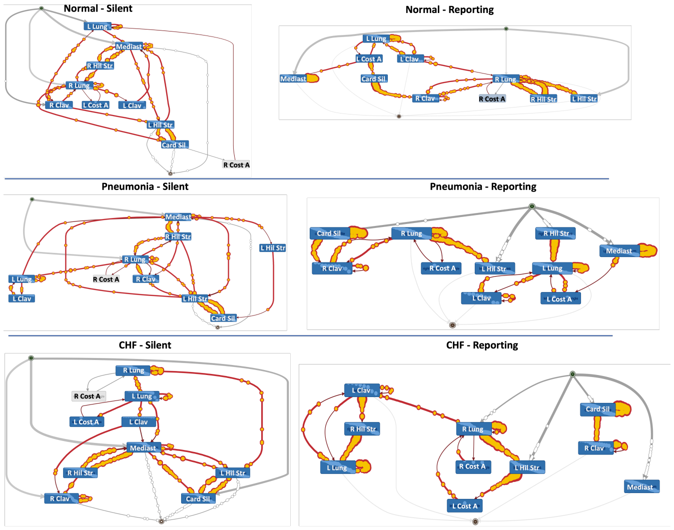

Figure 3: Markov chains representing the different anatomic regions that the radiologist visited during moments of silence vs. moments of reporting across different diagnoses. The yellow/red bubbles represent the radiologist's gaze flow when transitioning from one anatomical region to another.

obvious abnormalities. However, for the reporting phase, these figures vary slightly. For instance, it takes 12*sec* for the radiologist to report a Normal image, 20*sec* for a Pneumonia image, and 21*sec* for a CHF image. This increase in time is also in accordance with the Foveal verification phase of the 4-stage model, where the radiologist focuses on the perturbations to report any possible abnormalities.

In terms of the computed MCs, the initial observation of silent exploration reveals a more profound examination, as evidenced by the longer paths we observed. Analysis of MCs derived from the *Silent* stages demonstrates exploratory patterns, as the probability of accessing various anatomical regions from the initial state is consistent across different diagnoses. This is reflected in the resulting graphs, which exhibit higher depths, than those of *Reporting*, as all CXR regions must be traversed regardless of the starting point. On the other hand, an examination of Markov chains extracted from the Reporting stages reveals a targeted gaze towards specific anatomical regions. This is reflected in the resulting graphs, which exhibit a more horizontal pattern, as the radiologist's focus is directed towards regions of interest or regions potentially containing abnormalities.

4.3 Longest Repeating Visitation Patterns

The final results of this study correspond to the radiologist's frequent visitation pattern analysis. Our results are illustrated in Figure 4.

Our analysis of gaze patterns during the silent phase of the radiologist's image interpretation revealed that the frequency of gaze patterns is lower than in the reporting phase. Specifically, the radiologist looks back at the same regions fewer times. This finding is consistent with our previous analysis, which indicated that the silent phase is characterized by an exploratory pattern, where the

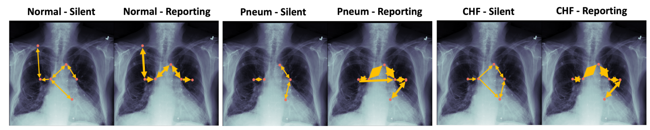

Figure 4: Most frequent visitation patterns using the longest repeating subsequence (LRS) algorithm. Thicker arrows depict more frequent patterns. Note that background CXR images are merely provided for context.

radiologist is more likely to examine a variety of regions within the image. These results suggest that the radiologist's gaze patterns during the silent phase are characterized by a more comprehensive exploration of the image rather than a focused examination of specific regions.

Regarding Reporting moments, our analysis of gaze patterns suggests that the frequency of eye fixations is significantly higher than in other phases. Specifically, the radiologist's eye gaze revisits the same regions repeatedly. This finding suggests that the radiologist focuses on specific regions within the image rather than a more exhaustive exploration during the reporting phase. This is consistent with exploitation patterns, where the radiologist focuses on describing regions of interest or potentially containing abnormalities. These results indicate that the radiologist's gaze patterns during the reporting phase are characterized by a more focused examination of specific regions rather than a broader exploration of the image.

5 CONCLUSIONS AND FUTURE WORK

The results of this study show that periods of silence indicate short and more exploratory gaze patterns which exhibit higher depths than those of Reporting phase, as all CXR regions must be traversed regardless of the starting point. When radiologists initially view chest X-ray images in silence, they apply their own specific and individual search patterns regardless of the type of disease present in the image. This indicates that the radiologists' search behaviors can be identified at this stage. However, when radiologists report their readings of the X-rays, the gaze patterns appear to be noisy and arbitrary.

In the future, we plan to continue our research by comparing the gaze patterns during the silent and reporting phases across different diagnoses using density map analysis. We will also investigate the Markov chains we have computed by analyzing their stationary distributions and transition matrices. This will allow us to gain a deeper understanding of the gaze patterns observed during these phases and potentially identify patterns that are specific to certain diagnoses.

In conclusion, this paper has argued that conventional approaches to deep learning applied to chest X-ray images have not been entirely successful. Despite the significant computational resources invested, simple function fitting has not provided conclusive results. We have therefore proposed a shift towards feature engineering, utilizing computer vision techniques to extract more meaningful semantics from fixation patterns. Furthermore, we have emphasized the importance of considering the unique characteristics of human-generated data, such as eye-gaze, and processing it in a manner that captures human functionality. While significant work remains to be done, we are optimistic that our trailblazing efforts will lead to promising advancements in the field.

ACKNOWLEDGMENTS

This material is based upon work supported by the UNESCO Chair on AI&XR; and the Portuguese *Fundação para a Ciência e a Tecnologia (FCT)* under grants no. 2022.09212.PTDC (XAVIER) and no. UIDB/50021/2020.

REFERENCES

Folami T Alamudun, Hong-Jun Yoon, Kathy Hudson, Garnetta Morin-Ducote, and Georgia Tourassi. 2015. Fractal analysis of radiologists' visual scanning pattern in screening mammography. In *Medical Imaging 2015: Image Perception, Observer Performance, and Technology Assessment* (Orlando, Florida, United States), Claudia R Mello-Thoms and Matthew A Kupinski (Eds.). SPIE.

Gary Aston-Jones and Jonathan D Cohen. 2005. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience* 28 (2005), 403–50.

K S Berbaum, E A Franken, Jr, D D Dorfman, E M Miller, E A Krupinski, K Kreinbring, R T Caldwell, and C H Lu. 1996. Cause of satisfaction of search effects in contrast studies of the abdomen. *Acad. Radiol.* 3, 10 (Oct. 1996), 815–826.

K S Berbaum, E A Franken, Jr, D D Dorfman, S A Rooholamini, C E Coffman, S H Cornell, A H Cragg, J R Galvin, H Honda, and S C Kao. 1991. Time course of satisfaction of search. *Invest. Radiol.* 26, 7 (July 1991), 640–648.

K S Berbaum, E A Franken, Jr, D D Dorfman, S A Rooholamini, M H Kathol, T J Barloon, F M Behlke, Y Sato, C H Lu, and G Y el Khoury. 1990. Satisfaction of search in diagnostic radiology. *Invest. Radiol.* 25, 2 (Feb. 1990), 133–140.

Ricardo Bigolin Lanfredi, Mingyuan Zhang, William F Auffermann, Jessica Chan, Phuong-Anh T Duong, Vivek Srikumar, Trafton Drew, Joyce D Schroeder, and Tolga Tasdizen. 2022. REFLACX, a dataset of reports and eye-tracking data for localization of abnormalities in chest x-rays. *Scientific Data* 9, 1 (2022), 1–15.

Tad T Brunyé, Marianna D Eddy, Ezgi Mercan, Kimberly H Allison, Donald L Weaver, and Joann G Elmore. 2016. Pupil diameter changes reflect the difficulty and diagnostic accuracy during medical image interpretation. *BMC Medical Informatics and Decision Making* 16 (2016).

D P Carmody, H L Kundel, and L C Toto. 1984. Comparison scans while reading chest images. Taught, but not practiced. *Invest. Radiol.* 19, 5 (Sept. 1984), 462–466.

Aarthi Ganesan, Maram Alakhras, Patrick C Brennan, and Claudia Mello-Thoms. 2018. A review of factors influencing radiologists' visual search behaviour. *Journal of medical imaging and radiation oncology* 62, 6 (2018), 747–757.

Mark S. Gilzenrat, Sander Nieuwenhuis, Marieke Jepma, and Jonathan D. Cohen. 2010. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, and Behavioral Neuroscience* 10 (2010), 252–69.

Dan Gusfield. 1997. *Algorithms on Strings, Trees, and Sequences.* Cambridge University Press.

Katarzyna Harezlak and Pawel Kasprowski. 2018. Application of eye tracking in medicine: A survey, research issues and challenges. *Computerized Medical Imaging and Graphics* 65 (2018), 176–190. https://doi.org/10.1016/j.compmedimag.2017.04.006 Advances in Biomedical Image Processing.

D. S. Hirschberg. 1975. A linear space algorithm for computing maximal common subsequence. *Commun. ACM* 18 (1975), 341—-343.

Hedvig Hricak, May Abdel-Wahab, Rifat Atun, Miriam Mikhail Lette, Diana Paez, James A Brink, Lluís Donoso-Bach, Guy Frija, Monika Hierath, Ola Holmberg, Pek-Lan Khong, Jason S Lewis, Geraldine McGinty, Wim J G Oyen, Lawrence N Shulman, Zachary J Ward, and Andrew M Scott. 2021. Medical imaging and nuclear medicine: a Lancet Oncology Commission. *The Lancet Oncology* 22, 4 (2021), e136–e172. https://doi.org/10.1016/S1470-2045(20)30751-8

Alistair E. W. Johnson, Tom Pollard, Roger Mark, Seth Berkowitz, and Steven Horng. 2019b. The MIMIC-CXR Database. https://doi.org/10.13026/C2JT1Q

Alistair E. W. Johnson, Tom J. Pollard, Seth J. Berkowitz, Nathaniel R. Greenbaum, Matthew P. Lungren, Chih-ying Deng, Roger G. Mark, and Steven Horng. 2019a. MIMIC-CXR-JPG: A large publicly available database of labeled chest radiographs. *CoRR* (2019). arXiv:1901.07042

Alexandros Karargyris, Satyananda Kashyap, Ismini Lourentzou, Joy T Wu, Arjun Sharma, Matthew Tong, Shafiq Abedin, David Beymer, Vandana Mukherjee, Elizabeth A Krupinski, and Mehdi Moradi. 2021. Creation and validation of a chest X-ray dataset with eye-tracking and report dictation for AI development. *Scientific Data* 8, 1 (March 2021), 92.

Harold L. Kundel and Calvin F. Nodine. 1975. Interpreting Chest Radiographs without Visual Search. *Radiology* 116 (1975), 527–532.

H L Kundel, C F Nodine, and E A Krupinski. 1990. Computer-displayed eye position as a visual aid to pulmonary nodule interpretation. *Invest. Radiol.* 25, 8 (Aug. 1990), 890–896.

Harold L. Kundel, Calvin F. Nodine, and Lawrence Toto. 1991. Searching for lung nodules: The guidance of visual scanning. *Investigative Radiology* 26 (1991), 777–781.

Damien Litchfield, Linden Ball, Tim Donovan, David J. Manning, and Trever Crawford. 2010. Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *Journal of Experimental Psychology: Applied* 16 (2010), 251–262.

Calvin F. Nodine and Harold L. Kundel. 1987. Using eye movements to study visual search and to improve tumor detection. *Radiographics* 7 (1987), 1241–50.

S Samuel, H L Kundel, C F Nodine, and L C Toto. 1995. Mechanism of satisfaction of search: eye position recordings in the reading of chest radiographs. *Radiology* 194, 3 (March 1995), 895–902.

Maurício Sousa, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge, and Daniel Simões Lopes. 2017. VRRRRoom: Virtual Reality for Radiologists in the Reading Room. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4057–4062. https://doi.org/10.1145/3025453.3025566

E L Thomas and E L Lansdown. 1963. Visual search patterns of radiologists in training. *Radiology* 81, 2 (Aug. 1963), 288–291.

W J Tuddenham. 1962. Visual search, image organization, and reader error in roentgen diagnosis. Studies of the psycho-physiology of roentgen image perception. *Radiology* 78, 5 (May 1962), 694–704.

William J. Tuddenham and William P. Calvert. 1961. Visual Search Patterns in Roentgen Diagnosis. *Radiology* 76, 2 (1961), 255–256. https://doi.org/10.1148/76.2.255 PMID: 13778547.

Gianpaolo Zammarchi and Claudio Conversano. 2021. Application of eye tracking technology in medicine: A bibliometric analysis. *Vision (Basel)* 5, 4 (Nov. 2021), 56.