# Fractal Analysis of Radiologists' Visual Scanning Pattern in Screening Mammography

Folami T Alamudun[a], Hong-Jun Yoon[c], Kathy Hudson[b], Garnetta Morin-Ducotte[b], Georgia Tourassi[c]

[a] Department of Computer Science and Engineering, Texas A&M University, College Station, TX, USA 77843;
[b] Department of Radiology, University of Tennessee Medical Center at Knoxville, Knoxville, TN, USA 37920;
[c] Health Data Sciences Institute, Biomedical Science and Engineering Center, Oak Ridge National Laboratory, One Bethel Valley Road, Oak Ridge, TN, USA 37831

## ABSTRACT

Several researchers have investigated radiologists' visual scanning patterns with respect to features such as total time examining a case, time to initially hit true lesions, number of hits, etc. The purpose of this study was to examine the complexity of the radiologists' visual scanning pattern when viewing 4-view mammographic cases, as they typically do in clinical practice. Gaze data were collected from 10 readers (3 breast imaging experts and 7 radiology residents) while reviewing 100 screening mammograms (24 normal, 26 benign, 50 malignant). The radiologists' scanpaths across the 4 mammographic views were mapped to a single 2-D image plane. Then, fractal analysis was applied on the composite 4-view scanpaths. For each case, the complexity of each radiologist's scanpath was measured using fractal dimension estimated with the box counting method. The association between the fractal dimension of the radiologists' visual scanpath, case pathology, case density, and radiologist experience was evaluated using fixed effects ANOVA. ANOVA showed that the complexity of the radiologists' visual search pattern in screening mammography is dependent on case specific attributes (breast parenchyma density and case pathology) as well as on reader attributes, namely experience level. Visual scanning patterns are significantly different for benign and malignant cases than for normal cases. There is also substantial inter-observer variability which cannot be explained only by experience level.

**Keywords:** visual perception, fractal analysis, mammography, gaze complexity, user modeling

## 1. INTRODUCTION

Breast cancer is the second leading form of cancer-related death affecting a large percentage of the female population. The mortality rate for this disease is largely dependent on early detection through the mammographic screening process [1]. However, studies show that the mammographic screening process is susceptible to different types of error resulting in misdiagnosis, with 50% of misdiagnosis resulting from human visual error [2–5]. Diagnostic error has received a lot of attention in recent years as members of the medical research community have focused on the visual processing and cognitive processes related to diagnostic decision making to better understand the causes of diagnostic error. In radiology, diagnostic errors can be attributed to visual search and recognition errors [6,7].

For over half a century, a large number of studies have focused on the radiologists' visual scan pattern during the image reading process. Findings from these studies indicate prevalence of errors in two broad areas: (1) how radiologists find what they are looking for (visual search); and (2) how radiologists interpret what they are looking at (image interpretation) [8–13]. A large body of eye-tracking research has also focused on gaining a better understanding of the relationship between visual search and diagnostic decision by analyzing radiologists' eye movements recorded during the diagnostic process [14–20].

Resent research work has shown the efficacy of eye-tracking in diagnostic performance prediction [21]. Voisin et al. conducted laboratory studies and applied machine learning techniques to predict error during the diagnostic characterization of mammographic lesions by combining features from radiologists' gaze behavior, and textural image characteristics[21]. Tourassi et al. investigated the relationship between radiologists' gaze, diagnostic decision, and image content of mammograms during mammographic cancer screening [22]. Their results show that machine learning can be used to build user dependent models to predict radiologists' medical errors by combining image content and gaze characteristics.

In mammography, all gaze tracking studies have been based on single view mammograms. This is not consistent with clinical practice where radiologists navigate simultaneously 4 different mammographic views. Furthermore, observed visual search has been typically summarized using features such as total time examining a case, time to initial hit on true lesions, total dwell time assessing a specific lesion, number of hits, etc. Although informative, these features fail to capture the gaze path trajectory. The purpose of this study is two-fold. First, we aim to examine and model the complexity of the radiologists' visual scanning pattern when viewing 4-view mammographic cases, as they typically do in clinical practice. Second, we aimed to understand if and how the radiologists' visual scanning pattern complexity is affected by 3 factors: (i) breast parenchyma density, (ii) case pathology, and (iii) radiologists' experience level.

## 2. METHODS

### 2.1. Experimental Protocol

To perform this study, ten readers of variable experience levels were recruited to conduct blind review of 100 four-view screening mammograms of varying pathology (see Table 1). Three of the 10 study participants were experienced MQSA-certified breast imagers while the remaining seven participants were radiology residents with at least one rotation in mammography (see Table 2).

**Table 1. Diagnostic cases**

| Type | No. of cases |
|---|---|
| Normal | 24 |
| Benign | 26 |
| Malignant | 50 |
| *Total* | *100* |

**Table 2. Participating radiologists characteristics**

| Type | Experience | No. of participants |
|---|---|---|
| Radiologist | > 10 yrs of practice | 2 |
| Radiologist | < 10 yrs of practice | 1 |
| Resident | > 2 mammo rotations | 4 |
| Resident | ≤ 2 mammo rotations | 3 |
| *Total* | | *10* |

Each reader was required to view 100 screen-film mammograms selected from the Digital Database of Screening Mammography (DDSM) [23]. Of the DDSM case corpus used for this study 24 cases were normal, 26 included biopsy-proven benign masses, and the 50 remaining cases included biopsy-proven malignant masses. The cancer cases did not include any benign lesions. Also, none of the study cases included calcifications. The radiologists were asked to report the location of any suspicious masses and corresponding BI-RADS rating as typically done in clinical practice.

### 2.2. Eye-tracking Apparatus & Data Collection

A customized graphical user interface was developed in-house for study participants to view each mammographic case and record their findings. Two medical grade monitors were used (dual-head 5MP mammo-grade Totoku LCD monitors calibrated to the DICOM display standard). The four views were displayed at low resolution (2 views per monitor) to fit the screen. The GUI provided the functionality of zooming in/outs, panning, and magnifying glass for detailed viewing of the mammographic views.

During the reading session, each reader was outfitted with an H6 head-mounted eye-tracker, with a 60 Hz sampling rate, and eye-head integration from Applied Science Laboratories (ASL, Bedford, Massachusetts, USA). Readers were instructed to view each case until satisfied with the viewing phase. When the reader was ready to give a diagnostic opinion, the eye-tracking recording phase was halted until the reader completed and reported their findings. After completion, the reader was instructed to proceed with viewing the next case. Prior to the study, each reader was carefully calibrated using the 9-point calibration protocol provided by ASL.

### 2.3. Data Analysis

Since gaze data for each case was collected from 4 mammographic views spread across two monitors, the raw eye-position data collected per case were mapped to a single two-dimensional image plane, maintaining the initial order in which the four views were presented during the reading session (i.e., LCC, RCC, LMLO, RMLO). Fixations were computed using an algorithm developed by Nodine and Kundel [24], and scan paths were derived connecting the fixation data.

For each case and reader viewing the case, his visual search scan path is a highly complex *gaze network*. We used fractal analysis to capture and model the complexity of these networks across cases and readers.
Fractal dimension is a mathematical tool for objective measurement of complex structures or patterns that cannot be readily described and quantified by application of Euclidian geometry. The network formed by connecting fixations during a mammographic screening can be treated as a fractal pattern, the dimension (***D***) of which, for two-dimensional objects, is expressed by a non-integer between 1 and 2. We used the *Minkowski–Bouligand box-counting method* [25] to estimate the fractal dimension (***D***) of each reader's gaze network for each examined case.

The interaction between gaze complexity, case pathology, case density, and radiologist's experience was evaluated using a four-factor fixed-effects ANOVA with 3 levels for case pathology (normal, benign, & malignant), 4 levels for case

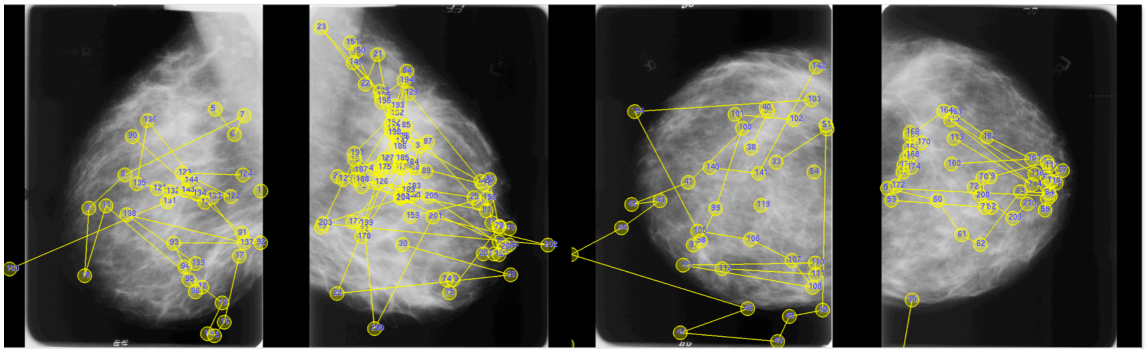

**Figure 1. Example image showing the initial ordering (LCC RCC LMLO RMLO)**

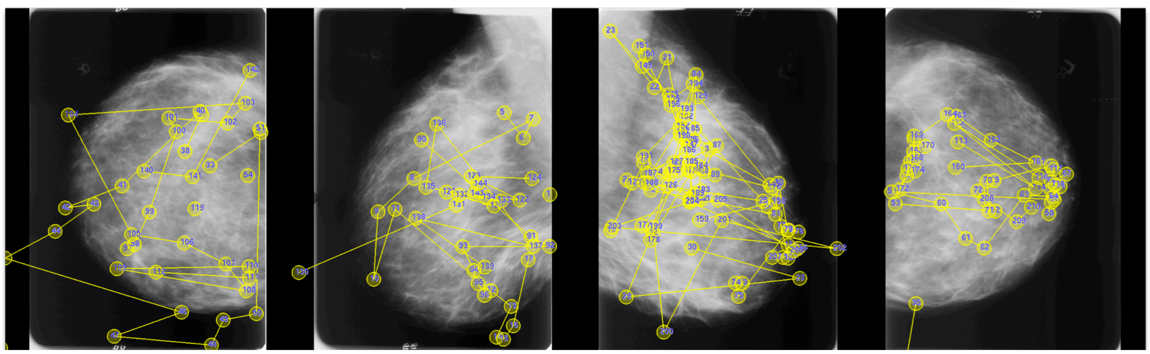

**Figure 2. Example image showing an alternate ordering (RCC RMLO LMLO LCC)**

density (fatty, fibroglandular, heterogeneous, dense), 3 levels for experience levels (new trainee, advanced trainee, expert), and 10 individual readers. We repeated the same data analysis process (computing complexity using fractal analysis followed by ANOVA) while changing the order in which the gaze data were mapped on a two-dimensional image plane (RCC, RMLO, LMLO, LCC).

## 3. RESULTS

We observed that *gaze networks* generated during mammographic screening were complex and appeared to be unique to each case and each reader. However, using ANOVA analysis on the estimated fractal dimensions we aimed to determine if these networks show any dependency with case pathology, breast density or the reader's experience level.

The interaction between gaze complexity, case pathology, case density, and readers's experience was evaluated using a four-factor fixed-effects ANOVA with 3 levels for case pathology (normal, benign, & malignant), 4 levels for case density (fatty, fibroglandular, heterogeneous, dense), 3 levels for experience levels (new trainee, advanced trainee, expert), and 10 individual readers. The results of the ANOVA tests for the two alternate mammographic view orderings are summarized in Table 3.

ANOVA showed that all four factors are independent predictors of the radiologists' visual scanning pattern complexity. This finding was consistent across both mammographic view orderings. None of the higher order effects were found to be significant.

**Table 3. Multi-factor ANOVA for both Fractal Dimensions of Initial Ordering and Alternate Ordering**

| Source | DoF | Initial Ordering | | Alternate Ordering | |
|---|---|---|---|---|---|
| | | F | p > F | F | p > F |
| Pathology | 2 | 13.53 | 1.62e-6 | 16.23 | 1.19e-7 |
| Density | 2 | 6.00 | 0.0026 | 3.91 | 0.0203 |
| Experience | 2 | 9.59 | 7.53e-5 | 12.89 | 3.00e-6 |
| Individual | 7 | 47.91 | < 1e-15 | 63.59 | < 1e-15 |
| Pathology : Density | 4 | 1.50 | 0.1998 | 1.34 | 0.2519 |
| Pathology : Experience | 4 | 0.93 | 0.4429 | 1.34 | 0.2515 |
| Density : Experience | 4 | 0.78 | 0.5349 | 0.74 | 0.5626 |
| Pathology : Individual | 14 | 1.64 | 0.0637 | 1.32 | 0.1865 |
| Density : Individual | 14 | 0.44 | 0.9616 | 0.64 | 0.8366 |
| Pathology : Density : Experience | 8 | 0.59 | 0.7860 | 0.63 | 0.7502 |
| Pathology : Density : Individual | 28 | 0.66 | 0.9110 | 0.51 | 0.9847 |
| Total | | 999 | | | |

Post-ANOVA t-tests with Bonferroni p-value adjustment were also performed (Table 4). Overall, the complexity of the readers' visual search patterns were significantly different between normal cases and cases including mass lesions. However, the malignancy status of the lesions did not affect the complexity of the visual search pattern. In addition, gaze pattern complexity was found to be significantly different between fibroglandular and heterogeneous/dense mammograms. However, this finding was not consistent between the two view orderings. Finally, the gaze complexity pattern was found to be significantly different between new trainees and experts for both orderings. There were some significant differences between other experience level groups, however these differences were not consistently significant between the two orderings.

**Table 4. Pairwise Comparisons of Groups of Case Pathology, Breast Density, and Radiologists' Experience Level**

| Pair 1 | Pair 2 | p-value Initial Ordering | p-value Alternate Ordering |
|---|---|---|---|
| Pathology – Normal | Pathology - Benign | 0.022 | 0.004 |
| Pathology – Normal | Pathology - Malignant | 2.4e-5 | 9.3e-6 |
| Pathology – Benign | Pathology - Malignant | 0.431 | 0.905 |
| Density – Fatty | Density - Fibroglandular | 0.604 | 0.600 |
| Density - Fatty | Density - Heterogeneous/Dense | 0.842 | 1.000 |
| Density - Fibroglandular | Density - Heterogeneous/Dense | 0.027 | 0.230 |
| Experience - New Trainee | Experience - Advanced Trainee | 0.291 | 1.9e-4 |
| Experience - New Trainee | Experience - Expert | 0.001 | 0.001 |
| Experience - Advanced Trainee | Experience - Expert | 0.025 | 1.000 |

Finally, statistical tests were performed to study the pairwise differences among the 10 readers (Table 5). Several significant pairwise differences were found suggesting that there was substantial inter-reader variability.

**Table 5. Pairwise Comparisons of Individual Readers (a) of Initial Ordering, and (b) of Alternate Ordering (N: New Trainee, A: Advanced Trainee, E: Expert).**

(a)

| | **N1** | **N2** | **A1** | **A2** | **A3** | **A4** | **A5** | **E1** | **E2** |
|---|---|---|---|---|---|---|---|---|---|
| **N2** | <1e-3 | | | | | | | | |
| **A1** | 1.000 | <1e-3 | | | | | | | |
| **A2** | 0.216 | 0.212 | 0.096 | | | | | | |
| **A3** | <1e-3 | 1.000 | <1e-3 | 0.032 | | | | | |
| **A4** | 1.000 | <1e-3 | 1.000 | 0.909 | <1e-3 | | | | |
| **A5** | 1.000 | <1e-3 | 1.000 | 0.014 | <1e-3 | 1.000 | | | |
| **E1** | <1e-3 | 1.000 | <1e-3 | 0.001 | 1.000 | <1e-3 | <1e-3 | | |
| **E2** | 1.000 | <1e-3 | 1.000 | 1.000 | <1e-3 | 1.000 | 1.000 | <1e-3 | |
| **E3** | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 |

(b)

| | **N1** | **N2** | **A1** | **A2** | **A3** | **A4** | **A5** | **E1** | **E2** |
|---|---|---|---|---|---|---|---|---|---|
| **N2** | <1e-3 | | | | | | | | |
| **A1** | 1.000 | <1e-3 | | | | | | | |
| **A2** | 1.000 | <1e-3 | 1.000 | | | | | | |
| **A3** | <1e-3 | 1.000 | <1e-3 | <1e-3 | | | | | |
| **A4** | 0.006 | <1e-3 | 0.001 | 0.002 | <1e-3 | | | | |
| **A5** | 0.073 | <1e-3 | 0.012 | 0.029 | <1e-3 | 1.000 | | | |
| **E1** | <1e-3 | 1.000 | <1e-3 | <1e-3 | 1.000 | <1e-3 | <1e-3 | | |
| **E2** | 1.000 | 0.014 | 1.000 | 1.000 | 0.002 | <1e-3 | <1e-3 | <1e-3 | |
| **E3** | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 | <1e-3 |

## 4. DISCUSSION

In this study, we analyzed the gaze networks formed by the radiologists' scan path during 4-view mammographic screening. We used fractal analysis to measure the complexity of these gaze networks. We also investigated two alternate ways of aggregating the gaze data collected from the 4 mammographic views to determine how the aggregation ordering impacts fractal analysis.

The findings of this study which were consistent across the two alternative 4-view gaze aggregation ordering were:

- Case pathology, case density, and radiologist's experience are all independent predictors of radiologists' visual search pattern complexity as measured by the fractal dimension.
- The average fractal dimension of the visual search patterns for normal cases is statistically significantly different from those of benign and malignant cases.
- The average fractal dimension of the visual search patterns of expert radiologists is statistically significantly different from that of new trainees.
- There are notable individual differences among radiologists.

In summary, our results demonstrate that the complexity of the visual search pattern in screening mammography measured by fractal analysis is dependent on case specific attributes (breast parenchyma density and case pathology) as well as on reader attributes, namely experience.

## 5. CONCLUSION

Visual search pattern complexity has been shown to have a significant dependency on case properties and radiologists' experience level. Given this observation, the next step of this investigation is to determine the efficacy of visual search pattern as a predictor of diagnostic error. If we can reliably predict the risk of diagnostic error using gaze complexity, we could leverage these predictive models to develop automated training and decision support systems personalized to the individual radiologist's needs.

## ACKNOWLEDGMENTS

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

## REFERENCES

[1] U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2011 Incidence and Mortality Web-based Report. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2014. Available at: www.cdc.gov/uscs.

[2] Bird, R. E., Wallace, T. W.., Yankaskas, B. C., "Analysis of cancers missed at screening mammography.," Radiology **184**(3), 613–617 (1992).

[3] Yankaskas, B. C., Schell, M. J., Bird, R. E.., Desrochers, D. A., "Reassessment of breast cancers missed during routine screening mammography: a community-based study," American Journal of Roentgenology **177**(3), 535–541, Am Roentgen Ray Soc (2001).

[4] Berlin, L., "Malpractice issues in radiology. Perceptual errors.," AJR. American journal of roentgenology **167**(3), 587–590, Am Roentgen Ray Soc (1996).

[5] Berlin, L., "Malpractice issues in radiology Errors in judgment," AJR-American Journal of Roentgenology **166**(6), 1259–1261, [Baltimore, MD, etc., American Roentgen Ray Society, etc.] (1996).

[6] Birdwell, R. L., Ikeda, D. M., O'Shaughnessy, K. F.., Sickles, E. A., "Mammographic Characteristics of 115 Missed Cancers Later Detected with Screening Mammography and the Potential Utility of Computer-aided Detection 1," Radiology **219**(1), 192–202, Radiological Society of North America (2001).

[7] Krupinski, E. A., "Current perspectives in medical image perception," Attention, Perception, & Psychophysics **72**(5), 1205–1217, Springer (2010).

[8] Krupinski, E. A., "Influence of experience on scanning strategies in mammography," Medical Imaging 1996, 95–101 (1996).

[9] Mello-Thoms, C., "How does the perception of a lesion influence visual search strategy in mammogram reading?," Academic radiology **13**(3), 275–288, Elsevier (2006).

[10] Mello-Thoms, C., "Visual search characteristics in mammogram reading: SFM vs. FFDM," SPIE Medical Imaging, 76270D–76270D (2010).

[11] Mello-Thoms, C., Dunn, S., Nodine, C. F., Kundel, H. L.., Weinstein, S. P., "The perception of breast cancer: what differentiates missed from reported cancers in mammography?," Academic radiology **9**(9), 1004–1012, Elsevier (2002).

[12] Mello-Thoms, C., Dunn, S. M., Nodine, C. F.., Kundel, H. L., "The perception of breast cancers-a spatial frequency analysis of what differentiates missed from reported cancers," Medical Imaging, IEEE Transactions on **22**(10), 1297–1306, IEEE (2003).

[13] Mugglestone, M. D., Gale, A. G., Cowley, H. C.., Wilson, A., “Defining the perceptual processes involved with mammographic diagnostic errors,” Medical Imaging 1996, 71–77 (1996).
[14] Donovan, T., Manning, D. J., Phillips, P. W., Higham, S.., Crawford, T., “The effect of feedback on performance in a fracture detection task,” Medical Imaging, 79–85 (2005).
[15] Kundel, H. L., “How to minimize perceptual error and maximize expertise in medical imaging,” Proc. SPIE **6515**, 651508–1 (2007).
[16] Kundel, H. L.., John Wright, D., “The Influence of Prior Knowledge on Visual Search Strategies During the Viewing of Chest Radioqraphs 1,” Radiology **93**(2), 315–320, The Radiological Society of North America (1969).
[17] Kundel, H. L.., La Follette Jr, P. S., “Visual Search Patterns and Experience with Radiological Images 1,” Radiology **103**(3), 523–528, The Radiological Society of North America (1972).
[18] Kundel, H. L., Nodine, C. F.., Carmody, D., “Visual scanning, pattern recognition and decision-making in pulmonary nodule detection,” Investigative radiology **13**(3), 175–181, LWW (1978).
[19] Kundel, H. L., Nodine, C. F.., Krupinski, E. A., “Searching for lung nodules: visual dwell indicates locations of false-positive and false-negative decisions,” Investigative radiology **24**(6), 472–478, LWW (1989).
[20] Manning, D., Ethell, S.., Donovan, T., “Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph,” British Institute of Radiology (2014).
[21] Voisin, S., Pinto, F., Morin-Ducote, G., Hudson, K. B.., Tourassi, G. D., “Predicting diagnostic error in radiology via eye-tracking and image analytics: Preliminary investigation in mammography,” Medical physics **40**(10), 101906, American Association of Physicists in Medicine (2013).
[22] Tourassi, G., Voisin, S., Paquit, V.., Krupinski, E., “Investigating the link between radiologists’ gaze, diagnostic decision, and image content,” Journal of the American Medical Informatics Association **20**(6), 1067–1075, BMJ Publishing Group Ltd (2013).
[23] Heath, M., Bowyer, K., Kegelmeyer Jr, D. K. P., Moore, R., Chang, K.., Munishkumaran, S., "Current status of the digital database for screening mammography." Digital mammography, 457-460. Springer Netherlands (1998).
[24] Nodine, C.., Kundel, H., “A visual dwell algorithm can aid search and recognition of missed lung nodules in chest radiographs,” Visual Search **2**, 399–406 (1990).
[25] Hausdorff, F., "Dimension und äußeres Maß." Mathematische Annalen 79, no. 1-2, 157-179. (1918).