Studying the Gaze Patterns of Expert Radiologists in Screening Mammography: A Case Study with Breast Test Wales

Lucie Lévêque
Laboratory Ergonomics Cognitive Sciences applied to Transport
Gustave Eiffel University
Lyon-Bron, France
lucie.leveque@ensc.fr

Philippa Young
Breast Test Wales
National Health Service
Cardiff, United Kingdom
philippa.young2@wales.nhs.uk

Hantao Liu
School of Computer Science and Informatics
Cardiff University
Cardiff, United Kingdom
liuh35@cardiff.ac.uk

***Abstract*—Eye-tracking technology has become a widely used means to understand how radiologists perceive and interpret medical images, providing useful information that can help improve diagnostic accuracy. However, existing eye-tracking studies in medical imaging remain limited due to the small number of stimuli and/or of subjects involved, and the lack of quantitative metrics to fully reveal readers' gaze behaviour. In this paper, we present the conduct of a larger scale eye-tracking study, where seven expert radiologists were asked to read 196 mammogram images. Furthermore, we carry out an analyse various gaze metrics including fixation duration, saccade amplitude, as well as gaze deployment, which quantify radiologists' gaze behaviour.**

Keywords—medical imaging, screening mammography, visual attention, eye-tracking, saliency

I. Introduction

According to the World Health Organisation (WHO), breast cancer is the most common invasive cancer for women, and the second main cause of cancer death after lung cancer [1]. Screening mammography uses low-dose X-ray imaging to detect early stage breast cancers, and so with a view to make treatment more successful for patients. However, lesion detection as conducted by human observers is prone to errors due to the inherent limitations of the human visual system (HVS). In radiology, estimates suggest that there may be up to a thirty percent miss rate and an equally high false-positive rate in some areas of the field [2]. It is therefore critical to better understand medical image perception and interpretation, and to use such knowledge to develop useful solutions to minimise errors in routine clinical practice [3]-[4].

The eye-tracking technology has been used in the literature to reveal how medical professionals read radiological images in order to improve their speed and accuracy in diagnostic reading, as well as to inform clinical radiology training programmes [5]-[6]. Advanced eye-tracking systems have indeed been developed to reliably track and record the eye movements of image readers. Through the analysis of eye-tracking data, useful information is provided that can help reveal visual search patterns and identify potential problems in radiological accuracy. In the literature, many eye-tracking studies have been undertaken in the field of radiology.

For instance, Kundel et al. [7] collected eye-tracking data from experienced mammographers, mammography fellows, and radiology residents who were asked to search for cancer in mammograms. The analysis focused on eye positions and attempted to find the percentage of cancer locations that were fixated within the initial viewing. The results suggested that the initial detection occurs before visual scanning takes place, as most of the cancer locations were fixated within the very first second of viewing. Voisin et al. [8] investigated the relationship between radiologists' gaze behaviour and diagnostic performance in the detection of lesions in mammogram images. The study recorded the eye movements of six readers who were asked to evaluate the likelihood of malignancy of forty cases of mammographic masses. By analysing various quantitative metrics derived from the eye-tracking data, such as the number of fixations, the fixation duration, and the fixation/saccade ratio, their study showed that these gaze behaviour characteristics were highly correlated with radiologists' diagnostic errors.

To improve the radiological reading practice, it is necessary to better understand how radiologists perform image search, detection and recognition tasks through the use of eye-tracking technology and gaze data analysis. In this paper, we present a new eye-tracking study in screening mammography with the following contributions.

- First, this study contains a large number of mammogram images with the aim to provide strong statistical evidence for the analysis of gaze behaviour. Note the existing studies often used a very limited number of stimuli.
- Second, rather than using the popular metrics, i.e., number of fixations and fixation duration to characterise the gaze behaviour, we propose to employ complementary information using the saccade amplitude.
- Third, we construct a gaze development representation, i.e., saliency map, to graphically quantify the gaze behaviour. This can also be used as a benchmark for the development of machine vision algorithms that can automatically predict radiologists' gaze patterns.
- Finally, the eye-tracking data will be made publicly available to the research community.

II. Eye-Tracking Experiment

A. Stimuli

The set of stimuli used in our eye-tracking experiment is composed of 196 mammogram images of the mediolateral oblique (MLO) view, one of the standard views in mammographic imaging acquisition. The MLO view presents the most important projection as it allows depicting most breast tissues, e.g., the representation of the pectoral muscle on the MLO view is a key component in assessing the adequacy of the mammogram. The mammograms were extracted from 98 anonymised cases (note all cases were known to be lesion-free with a view to encourage the participants to consider all plausible clinical outputs) from the University Hospitals KU Leuven in Belgium. The original resolution of the mammograms was either 2080×2800 pixels or 2800×3518 pixels. All the images were linearly downscaled to 1080×1920 pixels to enable a controlled experiment. Fig. 1 illustrates examples of images used in our experiment.

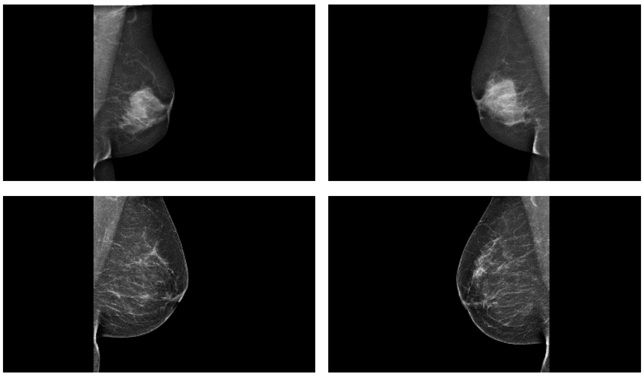

Fig. 1. Illustration of four sample stimuli from two patient cases (each case contains the MLO views of the left and right breasts) used in our eye-tracking experiment.

B. Experimental Procedure

The eye-tracking experiment was conducted in a mammography reading room at Breast Test Wales, Cardiff, United Kingdom. The venue represented a controlled viewing environment with low surface reflectance and constant ambient light. The viewing distance was approximately sixty centimetres. A 19-inch LCD monitor was used to display the stimuli and was calibrated to the Digital Imaging and Communications in Medicine (DICOM): Greyscale Standard Display Function (GSDF) [9]-[11].

Firstly, the MLO view of the left breast was displayed for three seconds. It was then replaced by the MLO view of the right breast of the same patient, also displayed for three seconds. Once a participant had viewed both images, they would be prompted to answer the question "refer or not refer" by looking at one of these two options on the screen. This was meant to simulate the procedure of routine breast screening practice. Indeed, in routine, suspicious cases require further investigation. Fig. 2 illustrates the sequence of viewing.

The eye movements of the observers were recorded using a SensoMotoric Instrument (SMI) Red-m advanced eye-tracker at a sampling rate of 250 Hz, using a nine-point calibration protocol.

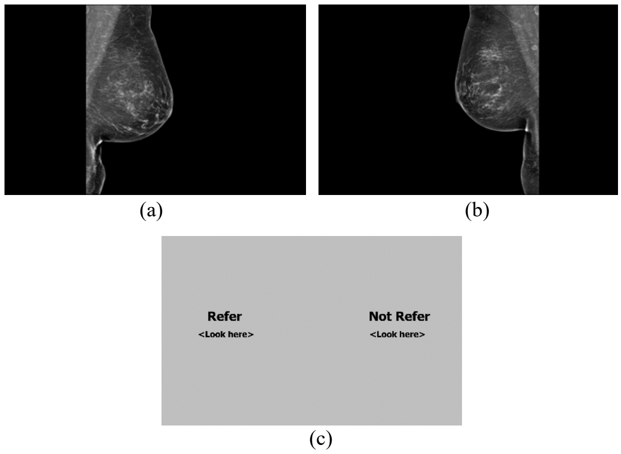

Fig. 2. Illustration of the experimental procedure: (a) MLO view of a left breast, (b) MLO view of the corresponding right breast, and (c) the question asked after viewing (a) and (b).

C. Participants

Breast Test Wales (BTW) [12] is a Welsh organisation delivering the National Health Service Breast Screening Programme (NHSBSP). Every three years, women aged fifty to seventy years old are offered a screening mammography exam by the programme. There are three regional centres in Wales, i.e., Cardiff, Swansea, and Llandudno. Each centre has an important clinical team composed of various occupations, including radiologists, surgeons, breast physicians, pathologists, etc.

Seven expert radiologists from the Cardiff centre, having different degrees of experience in mammography (i.e., ranging from two to twenty-five years of experience), participated in the experiment. The radiologists were referred to as R1, R2, R3, R4, R5, R6, and R7, having, respectively, two, five, six, eight, ten, twenty, and twenty-five years of experience in mammogram reading.

III. Experimental Results

Upon completion of the eye-tracking experiment, gaze information was extracted from the raw data using the SMI BeGaze Analysis software, including the number of fixations and saccades per image, their duration, and their coordinates.

A. Analysis of the Fixations

Fixations are the most common feature when analysing eye movements. A fixation has been defined by SMI's software using the dispersal and duration-based algorithm as established in [13], with the minimum fixation threshold being set to 100 ms. The fixation duration has been widely used as a quantitative metric to analyse the viewing behaviour of human subjects.

In radiology, readers with different degrees of experience (e.g., experts and novices) can be characterised by their average gaze/fixation duration [14]. Fig. 3 shows the mean duration of fixations over all stimuli used in our experiment for each of the seven radiologists. In this figure, two clusters can be clearly observed, a first cluster including R1, R2, R3, R4 and R5, and a second cluster including R6 and R7.

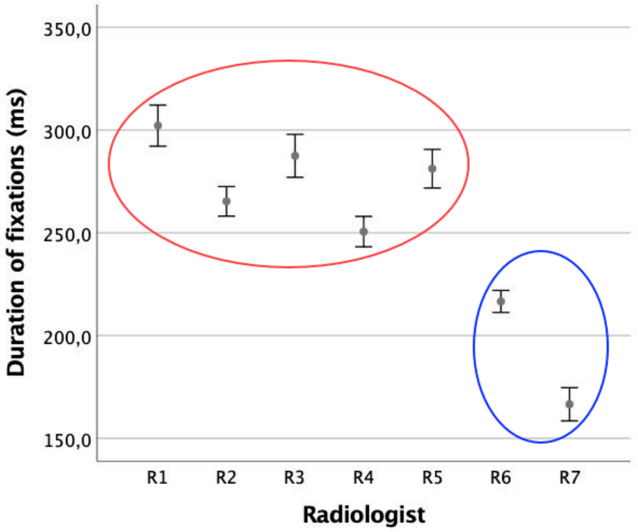

Fig. 3. Illustration of the mean fixation duration, over all fixations recorded for all test stimuli, for each radiologist. The cluster circled by red colour corresponds to "junior" radiologists, whereas the cluster circled by blue colour corresponds to "senior" radiologists. Error bars indicate a 95% confidence interval.

It can be also seen from Fig. 3 that "senior" radiologists, i.e., R6 and R7, having more than twenty years of experience in mammography, had shorter fixations than "junior" radiologists (i.e., ≤ 10 years of experience in mammography). The observed difference was further analysed using statistical hypothesis testing, i.e., an ANOVA (Analysis of Variance) was performed. The results of the statistical test show that there is a significant difference between the two groups of radiologists (i.e., p-value < 0.05). The results also show that there is a statistically significant difference (i.e., p-value < 0.05) between the two "senior" radiologists, i.e., R6 and R7. The radiologist who has twenty-five years of experience had the shortest fixation duration.

B. *Analysis of the Saccades*

Saccades can be defined as quick and simultaneous movements happening between fixations [15]. The saccade amplitude represents the Euclidian distance between two consecutive fixations, and is expressed in degree of visual angle [16]. Fig. 4 illustrates the distribution of saccade amplitude averaged over all stimuli for the senior radiologists on the one hand, and for the junior radiologists on the other.

One can notice from Fig. 4 a peak of saccade amplitude between one and three degrees of visual angle for both groups. Contrary to the senior radiologists, the junior present very few saccades with an amplitude bigger than ten degrees of visual angle. Concretely, this means that, in general, radiologists show very short trajectories between two fixated points and, therefore, seem to focus their gaze in a relatively small region of interest.

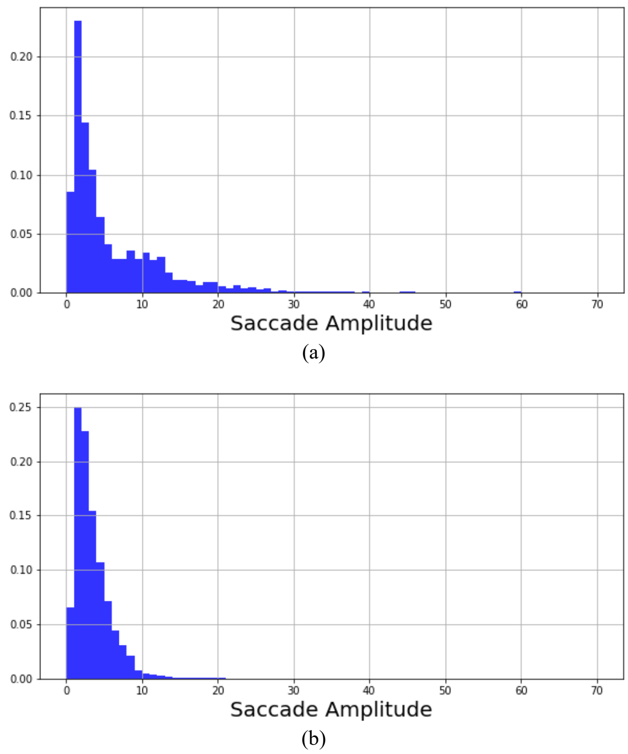

Fig. 4. Illustration of the distribution of the saccade amplitude over all stimuli used in our experiment for: (a) the senior radiologists, and (b) the junior radiologists.

C. *Analysis of the Gaze Deployment*

The collected eye-tracking data were also used to generate saliency maps, i.e., topographic representations indicating conspicuousness of scene location [17]-[19]. To create a saliency map, each fixation location gives rise to a greyscale patch simulating the foveal vision of the human visual system. The activity of a given patch is modelled as a Gaussian distribution, where the standard deviation approximates the size of the human fovea [20]. A saliency map *SM* can thus be calculated as follows:

$$SM_i(k,l) = \sum_{j=1}^{T} \exp\left(-\frac{(x_j-k)^2+(y_j-l)^2}{\sigma^2}\right) \quad (1)$$

where $SM_i(k,l)$ corresponds to the saliency map of the stimulus I_i (x_j,y_j) corresponds to the spatial coordinates of the j-th fixation ($j = 1\ldots T$), with T being the total number of fixations over all observers; and σ corresponds to the standard deviation of the Gaussian (i.e., two degrees of visual angle).

Fig. 5 illustrates the saliency maps (the brighter the regions, the higher the saliency) created from our eye-tracking data for two sample patient cases, for the "junior" and "senior" radiologists, respectively. It can be observed from the figure that the "senior" radiologists tend to produce a more concentrated saliency map than the "junior" radiologists, whose gaze is more dispersed. This suggests that there is a strong agreement between "senior" radiologists.

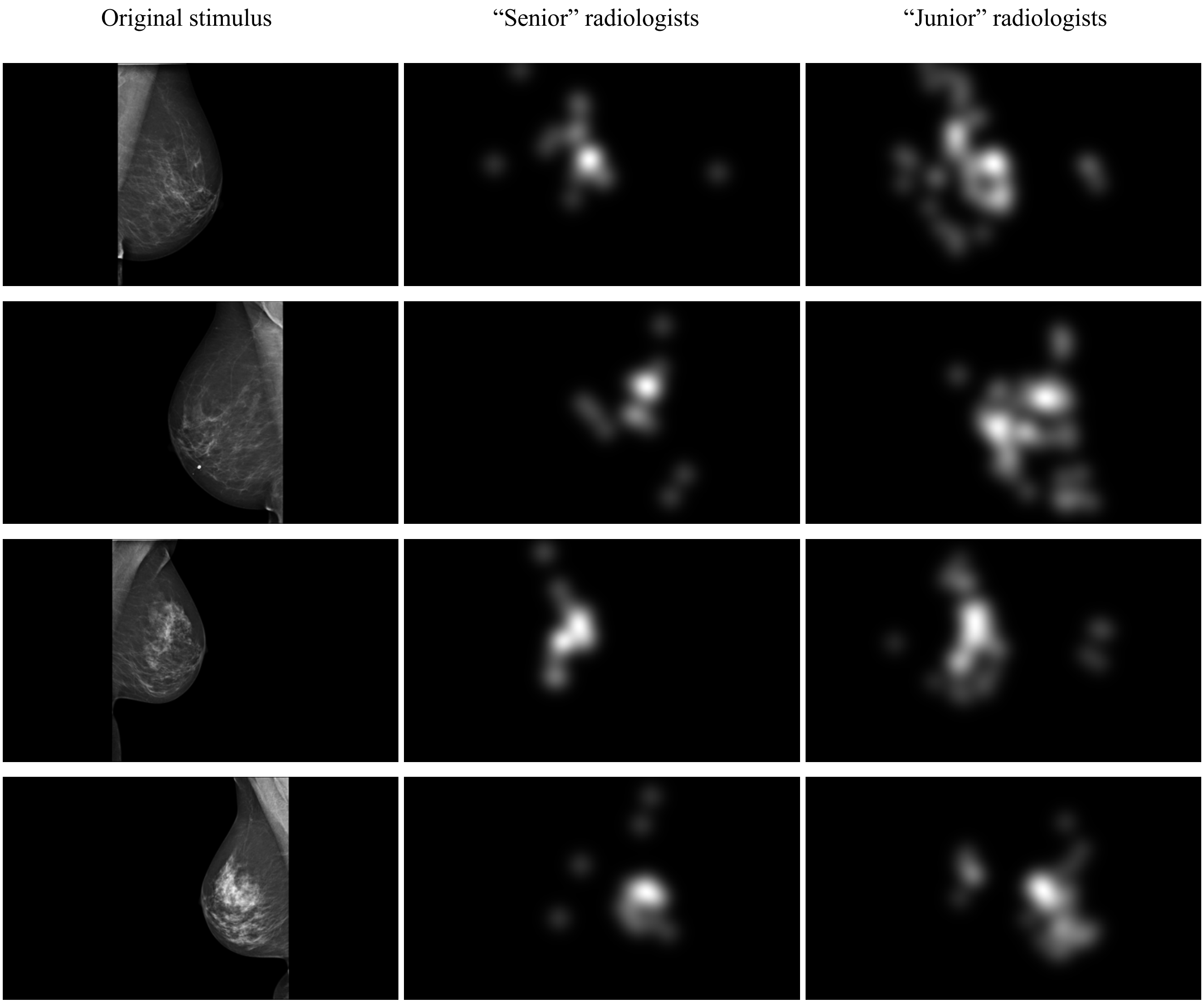

Fig. 5. Illustration of the saliency maps constructed for two patient cases for the “junior” and “senior” radiologists. The brighter the regions, the higher the saliency.

IV. Conclusions

In this paper, we carried out a large-scale eye-tracking experiment with mammogram images. Seven expert radiologists from Breast Test Wales participated in the study. A statistical analysis of fixations and saccades, as well as a study of gaze deployment, revealed that there is a significant difference in gaze behaviour between radiologists working in the same environment depending on their degrees of experience. The findings can inform further study on the development of tailored image processing algorithms to optimise individual performance.

References

[1] World Health Organisation, “International Agency for Research on Cancer”, *Press Release 263*, September 2018.

[2] C. Lee, P. Nagy, S. Weaver, and D. Newman-Toker, “Cognitive and System Factors Contributing to Diagnostic Errors in Radiology”, *American Journal of Roentgenology*, vol. 201, pp. 611-617, September 2013.

[3] E. Krupinski, “Visual Scanning Patterns of Radiologists Searching Mammograms”, *Academic Radiology*, vol. 3, pp. 137-144, February 1996.

[4] P. Agrawal, M. Vatsa, and R. Singh, “Saliency Based Mass Dectection from Screening Mammograms”, *Signal Processing*, vol. 99, pp. 29-47, June 2014.

[5] L. Lévêque, H. Bosmans, L. Cockmartin, and H. Liu, “State-of-the-Art: Eye-Tracking Studies in Medical Imaging”, *IEEE Access*, vol. 6, pp. 37023-37034, June 2018.

[6] L. Lévêque, B. Vande Berg, H. Bosmans, L. Cockmartin, M. Keupers, C. Van Ongeval, and H. Liu, “A Statistical Evaluation of Eye-Tracking Data on Screening Mammography: Effects of Expertise and Experience on Image Reading”, *Signal Processing: Image Communication*, vol. 78, pp. 86-93, October 2019.

[7] H. Kundel, C. Nodine, E. Krupinski, and C. Mello-Thoms, “Using Gaze-Tracking Data and Mixture Distribution Analysis to Support a Holistic Model for the Detection of Cancers on Mammograms”, *Academic Radiology*, vol. 15, pp. 881-886, July 2008.

[8] S. Voisin, F. Pinto, G. Morin-Ducote, K. Hudson, and G. Tourassi, “Predicting Diagnostic Error in Radiology via Eye-Tracking and Image Analytics: Preliminary Investigation in Mammography”, *Medical Physics*, vol. 40, October 2013.

[9] E. Samei, "Assessment of Display Performance for Medical Imaging Systems", *Report of the American Association of Physicists in Medicine, Task Group 18, Medical Physics Publishing*, April 2005.

[10] B. Hemminger, R. Johnston, J. Rolland, and K. Muller, "Introduction to Perceptual Linearization of Video Display Systems for Medical Image Presentation", *Journal of Digital Imaging*, vol. 8, pp. 21-34, February 1995.

[11] National Electrical Manufacturers Association, Digital Imaging and Communications in Medicine (DICOM) Part 14: Greyscale Standard Display Function, Rosslyn, United States of America, 2004. [Online]. Last accessed on January 2020.

[12] Breast Test Wales – Bron Brawf Cymru. [Online]. Last accessed January 2020.

[13] D. Salvucci, and J. Goldberg, "Identifying Fixations and Saccades in Eye-Tracking Protocols", *Proceedings of the Eye Tracking Research and Applications Symposium, Tampa, United States of America*, pp. 71-78, November 2000.

[14] E. Krupinski, "Visual Scanning Patterns of Radiologists Searching Mammograms", *Academic Radiology*, vol. 3, pp. 137-144, February 1996.

[15] O. Le Meur, and Z. Liu, "Saccadic Model of Eye Movements for Free-Viewing Condition", *Vision Research*, vol. 116, pp. 152-164, November 2015.

[16] O. Le Meur, A Coutrot, and Z. Liu, "Visual Attention Saccadic Models Learn to Emulate Gaze Patterns from Childhood to Adulthood", *IEEE Transactions on Image Processing*, vol. 26, pp. 4777-4789, June 2017.

[17] C. Privitera, and L. Stark, "Algorithms for Defining visual Regions-of-Interest: Comparison with Eye Fixations", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 970-982, September 2000.

[18] H. Liu, and I. Heynerickx, "Visual Attention Modeled with Luminance Only: From Eye-Tracking Data to Computational Models", *5th International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, United States of America*, January 2010.

[19] H. Liu, and I. Heynderickx, "Visual Attention in Objective Image Quality Assessment: Based on Eye-Tracking Data", *IEEE Transactions Circuits and Systems for Video Technologies*, vol. 21, pp. 971-982, July 2011.

[20] H. Liu, and I. Heynderickx, "Visual Attention in Objective Image Quality Assessment: Based on Eye-Tracking Data", *IEEE Transactions Circuits and Systems for Video Technologies*, vol. 21, pp. 971-982, July 2011.

[21] T. Judd, F. Durand, and A. Torralba, "A Benchmark of Computational Models of Saliency to Predict Human Fixations", *MIT Technical Report*, January 2012.

[22] J. Karel, C. Koch, and P. Perona, "Graph-Based Visual Saliency", *Proceedings of Advances of Neural Information Processing Systems, Pasadena, United States of America*, pp. 545-552, December 2006.

[23] L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 1254-1259, November 1998.

[24] N. Riche, M. Mancas, M. Duvinage, M. Mibulumukini, B. Gosselin, and T. Dutoit, "RARE2012: A Multi-Scale Rarity Based Saliency Detection With its Comparative Statistical Analysis", *Image Communication*, vol. 28, pp. 642-658, July 2013.