IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received February 17, 2020, accepted March 1, 2020, date of publication March 16, 2020, date of current version March 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2980901

ISeeColor: Method for Advanced Visual Analytics of Eye Tracking Data

KAREN PANETTA[1], (Fellow, IEEE), QIANWEN WAN[1], (Student Member, IEEE), SRIJITH RAJEEV[1], (Student Member, IEEE), ALEKSANDRA KASZOWSKA[2], AARON L. GARDONY[3], KEVIN NARANJO[1], HOLLY A. TAYLOR[2], AND SOS AGAIAN[4], (Fellow, IEEE)

[1]Department of Electrical and Computer Engineering, Tufts University, Medford, MA 02155, USA
[2]Department of Psychology, Tufts University, Medford, MA 02155, USA
[3]U.S. Army Combat Capabilities Development Command Soldier Center, Natick, MA 01760, USA
[4]Department of Computer Science, The City University of New York, New York, NY 10019, USA

Corresponding author: Srijith Rajeev (srijith.rajeev@tufts.edu)

This work was supported by the U.S. Army Combat Capabilities Development Command under Agreement W911QY-15-2-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Combat Capabilities Development Command, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

ABSTRACT Recent advances in head-mounted eye-tracking technology have allowed researchers to monitor eye movements during locomotion in real-world environments, increasing the ecological validity of research on human gaze behavior. While collecting eye-tracking data is becoming more accessible, visual analytics of eye-tracking data remains difficult and time-consuming. As such, there is a significant need for developing efficient visualization and analysis tools for large-scale eye-tracking data. This work develops a first-of-its-kind eye-tracking data visualization and analysis system that allows for automatic recognition of independent objects within field-of-vision, using deep-learning-based semantic segmentation. This system recolors the fixated objects-of-interest by integrating gaze fixation information with semantic maps. The system effectively allows researchers to automatically infer what objects users view and for how long in dynamic contexts. The contributions are 1) a data visualization and analysis system that uses deep-learning technology along with eye-tracking data to automatically recognize objects-of-interest from head-mounted eye-tracking video recordings, and 2) a graphical user interface that presents objects-of-interest annotation along with eye-tracking data information. The architecture is tested with an outdoor case study of users walking around the Tufts University campus as part of a navigation study, which was administered by a team of research psychologists.

INDEX TERMS Eye-trackers, cognitive science, data visualization, data analysis, deep-learning.

I. INTRODUCTION

Numerous research areas and commercial products utilize head-mounted eye-tracking devices, such as education [1], cognitive psychology [2], usability marketing [3], [4], on-road driving applications [5], medical applications [6], [7], information visualization research [8]–[10], eye-control accessibility, and assistive technology [10], [11].

Head-mounted eye-trackers are lightweight and unobtrusive, which enables the recording of eye movements without restricting movement in more naturalistic experimental settings [12]–[14]. High-frequency infrared eye cameras continuously monitor and record pupil position by tracking the distance between the center of the pupil and the position of artificially illuminated corneal reflections [15], [16]. Offline analysis of this data yields information about discrete eye movement events such as fixation, saccades, smooth pursuit, and scanpath.

Ultimately, present-day eye-tracking technology allows researchers to extract eye movement data to indicate where the participants looked, what they looked at, and for how long the viewing had occurred [17], [18]. Data visualization and analytic approaches have been developed to better understand the nature of eye movements [19]–[21]. However, visualizing and analyzing eye-tracking data collected during real-world locomotion poses a particular analytic challenge for automatically obtaining information on where a person is looking at and what a person is looking at in a specific scene [22]–[25].

The associate editor coordinating the review of this manuscript and approving it for publication was Feng Xia.

TABLE 1. Research questions that eye-tracking researchers are investigating.

Property	Definition
Q1: where	The position of the recorded raw gaze points in the scene video frame.
Q2: what	The semantic interpretation from the region-of-interest to object category within the visual field.
Q3: how long	The duration of eye movement events falling within a specific objects-of-interest within FOV.

Most eye–tracking analysis systems utilize manual annotation systems which are challenging and time-consuming. Few eye-tracking analysis systems exist that can enable the automatic categorization of what people looked at without manually tagging the areas-of-interest (AOI) in the video. However, there are no automated visual analytic tools that can present all three aspects (Q1: where, Q2: what, and Q3: how long) in a dynamic scene video. They are described in more detail in Table 1.

Table 2 describes the existing eye-tracking methodologies along with their descriptions and features (Q1, Q2, Q3). To the best of the author's knowledge, there are currently no solutions to automate visualization of massive head-mounted eye-tracking data that combines the Q1: where, Q2: what, and Q3: how long information types in a single visualization and analysis tool.

The presented ISeeColor software architecture tackles these three primary questions (Q1-Q3) for eye-tracking data visualization and analysis. It integrates gaze direction information (where) from the hardware, enables automatic recognition for fixated objects in areas-of-interest (objects-of-interest) using image semantic segmentation (what), and facilitates data visualization using fast-speed image recoloring based on the fixation duration (how long). The main contributions of ISeeColor are:

i ISeeColor can automatically annotate objects-of-interest (OOI) by adapting deep-learning-based scene semantic segmentation. This minimizes the need for time-consuming and error-prone manual annotation.
ii ISeeColor implements efficient and low-cost OOI recoloring. It uses color to represent fixation duration values. Color is a simple and intuitive indicator for data visualization and for highlighting certain OOI based on gaze density. Furthermore, the color scale can be set by the user.
iii ISeeColor offers a user-friendly graphical user interface (GUI), where OOI's are recolored while the color intensity changes based on the fixation duration overlaid on a video recording from the scene camera.

This paper is organized as follows: Section 2 presents related work. Section 3 shows the proposed architecture of ISeeColor and an illustrative "walking-around-campus" case study designed and conducted by a team of Tufts University research psychologists. Section 4 presents the evaluation and comparison of the proposed ISeeColor. Lastly, section 5 sum-

TABLE 2. Related work on eye-tracking data visualization and analysis approaches, Q1: where, Q2: what, and Q3: how long. ∗ means Listed related work attempted to answer only a subset of Q1,Q2, and Q3. ISeeColor is the only eye-tracking data visualization and analysis which automates Q1, Q2, and Q3.

Category	Method	Related work	Description	Feature
Tradition Statistics	Statistical metrics	[28]	Display eye movement information in a numeral format, such as fixation count, saccade/fixation ratio, saccade amplitudes.	Q1 Q3
Eye-tracking data visualization	Attention heatmap	[17] [33] [34] [35] [30] [36]	A visualization technique that can reflect the distribution of gaze density.	Q1 Q3
	Scanpath visualization	[29] [37] [38] [39]	A viewing path which incorporates fixations and saccades while a person is looking at a scene.	Q1 Q3
	AOI visualization techniques	[40]	Annotation of the observers' regions-of-interest on the stimulus.	Q1 Q2
Eye-tracking data visualization and analysis	Combination of DOI, AOI and eye-movements metrics	[6] [19] [20] [30] [22] [41] [42] [43] [32]	Diverse eye-tracking data visualization and analysis systems that provide data points, eye-movements, scan-paths, or heat-maps.	*(Q1 Q2 Q3)
	Combines Statistical, AOI, and eye-movements metrics along with deep-	**ISeeColor**	**Provides visualization and analysis of position, classification, and timeline of the AOI.**	**Q1 Q2 Q3**

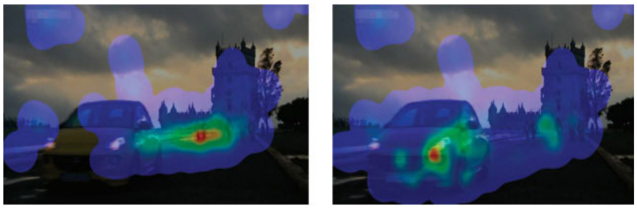

FIGURE 1. Attention maps for two images taken from the same video scene. In the left image, the area colored in red color indicates that the two people in the background gained the most fixation time. In the right image, the most attended to region is the moving car [30].

marizes the contributions and discusses the future directions of this work.

II. RELATED WORK

A. RELATED TERMINOLOGY

Eye-tracking technology allows us to infer which aspect of the visual scene a person is looking at, that is, where they are deploying their gaze. Gaze deployment, the direction of one's vision, is often described in conjunction with overt visual attention. Attention can be defined as the internal distribution of processing resources, and eye movements are referenced to mark the path of attention through a scene [26]. The correlation between eye movements and attention is used to study how individuals perform tasks like reading a book, driving a car, or exploring a scene. All these tasks require serial processing of different fragments of a scene based on eye movements [27].

Eye movements are broadly divided into four categories. They are:*(1) Fixations*: eye movements that stabilize the retina over a stationary OOI, (2)*Saccades*: rapid eye movements that occur in between fixations, *(3) Pursuit movements*: occur when a person is visually tracking a moving target, and (4)*Scanpath*: a sequence of saccades and fixations which constitute viewing.

B. EYE-TRACKING DATA VISUALIZATION AND ANALYSIS

Eye-tracking data is generally visualized with the help of visual overlays such as heatmaps and scanpaths. In heatmaps, a higher color intensity is indicative of longer fixation duration. In scanpaths, the sequence of fixations and saccades are denoted as a path traveled over a scene. An additional consideration in visualizing eye-tracking data is defining the areas-of-interest (AOI), which are objects within the visual scene that are of particular interest to researchers for analysis. The related literature for eye-tracking data visualization approaches and trends for developing data visualization and analysis systems are reviewed below:

1) CONVENTIONAL EYE-TRACKING DATA VISUALIZATION

The wide variety of eye movement metrics available to researchers require data visualization techniques to continuously evolve and accommodate advances in research. Eye movement metrics [28], such as fixation duration count, saccade/fixation ratio, and saccade amplitudes, can be easily computed from raw eye-tracking data using standard software.

Scanpath visualization incorporates fixations and saccades in creating a "viewing path" while a person is looking at a scene. In a typical scanpath visualization, each fixation is indicated by a circle, where the radius corresponds to the fixation duration. Saccades between fixations are represented by connecting lines between these circles [29] (See Figure 2).

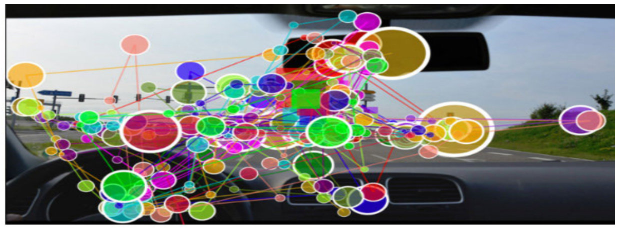

FIGURE 2. Eye movements of twenty drivers visualized as a scanpath in an autonomous driving test. In this image, larger circles indicate longer fixation durations [31]. Car manufacturers are developing systems to support drivers and warn them of things that may not have been perceived fast enough based on collected gaze information.

AOI (area-of-interest) based visualization techniques also consider the annotation of the objects on the stimulus that are of special interest to the researcher. Gaze Stripes [40] is an image-based visualization technique for eye-tracking data, which can display image data around the gaze points over time. However, AOI based eye-tracking methods do not define the area by the shape of the object.

2) LARGE-SCALE EYE-TRACKING DATA VISUALIZATION AND ANALYSIS

When conventional data visualization techniques, as those previously presented, no longer support analysis of massive eye-tracking datasets, the emerging discipline of data visualization and analysis can assist in explorative large-scale eye-tracking data analysis.

ISeeCube [20] uses space-time cube visualization [30] (STC) for eye-tracking data visual analysis. Kurzhals *et al.* created a timeline visualization to show AOI-based scanpaths of different viewers based on manual annotation of AOIs. Kurzhals *et al.* [22] described an AOI annotation process using automatic clustering of eye-tracking data integrated into an interactive labeling and analysis system (see Figure 3).

FIGURE 3. AOI based manual annotation and data visualization analysis approach for mobile eye-tracking [22].

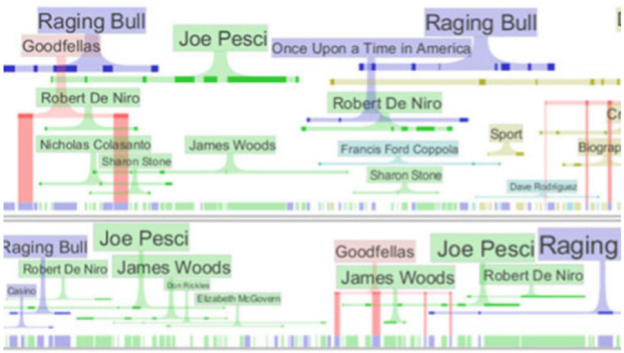

FIGURE 4. A scarf plot of a movie DOI data visual analytics [32]. They are scaled and recolored according to how much they were viewed around particular time-points. Note that DOIs are labeled manually.

eSeeTrack [19] was proposed for visualizing fixation patterns from video data collected by eye-tracking devices. It also enables users to manually label each fixation with its semantic content. GazeDx [6] was proposed as an interactive visual analytics tool for the comparative analysis of gaze data and spatial view of volumetric medical images provided by radiologists. A visual analytics approach [41] involving multiple concurrent evaluation procedures, such as ''thinking aloud'' protocols, interaction logs, and eye-tracking were also proposed.

Recently Muthumanickam et al [42] introduced a novel method to visualize spatial AOIs for long duration eye-tracking studies. It combines clustering and cluster merging changing over time. In addition, there is an increasing interest in methods for data-of-interest (DOI) eye-tracking analysis [43]. Jianu and Alam [32] established a foundation for the gaze to object mapping or DOI analysis by translating concrete data into visual representations, see Figure 4.

III. THE PROPOSED SYSTEM ARCHITECTURE FOR ISEECOLOR AND ILLUSTRATION

Our proposed system architecture ''ISeeColor'', a head-mounted eye-tracking data visualization and analysis tool is illustrated in Figure 5 and described in algorithm I.

A. DATA ACQUISITION

Data collection for this research was performed using SMI eye-tracking glasses with 0.5° gaze position accuracy over all distances [44]. Calibration was performed using a 3-point calibration mode with a gaze cursor. Real-world eye-tracking data was collected as part of an outdoor case study of users walking around the Tufts University campus. A team of research psychologists captured the eye-tracking videos.

Let the video captured through the front scene camera of the eye-tracker be denoted by V_0, and the eye-gaze fixation data obtained from the two infrared cameras that monitor pupil movement be denoted by D_0. Let the fixation locations in each frame be denoted as $F_1^l F_2^l, \ldots F_k^l F_1^l$, $F_2^l, \ldots F_k^1$. Note that, additional interpolation between the generated eye movement information and recorded front

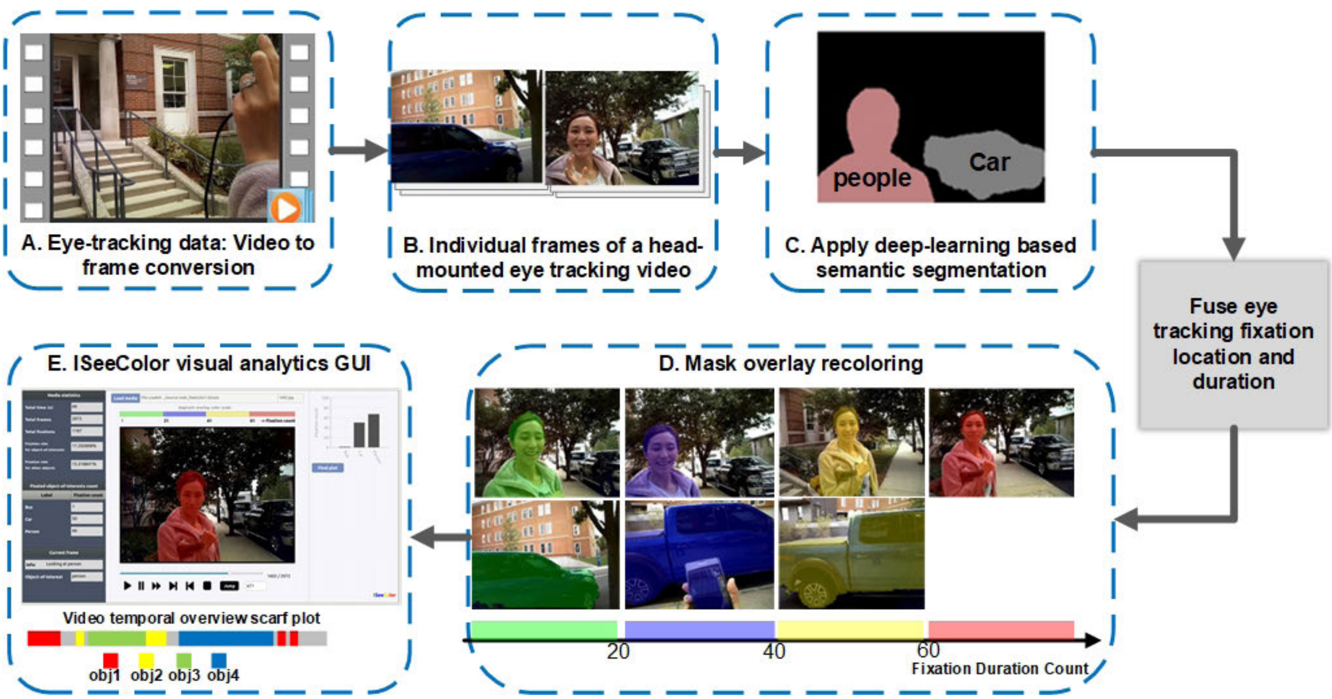

FIGURE 5. Overview of ISeeColor architecture demonstrated on an outdoor case study investigating human navigation. The input (A) mobile eye-tracking front scene video data is first broken down into (B) individual image frames. Then a deep-learning-based image semantic segmentation is implemented to (C) segment the pre-defined OOI in every frame. This is followed by (C) using the correlation of fixation coordinates along with segments in a frame to identify the fixated OOI, and (D) ISeeColor recolors fixated OOI while the color intensity changes as fixation duration increases. Note that, an abrupt color changes once the fixation duration value meets the pre-defined threshold. Finally, (E) information including where, what, and how long presents in the ISeeColor data visualization and analysis GUI.

Algorithm 1 ISeeColor Architecture

Input: head-mounted eye-tracking video: V_0;
eye movements data: D_0;

Initialization: individual image frames $I_1 I_2, \ldots I_k$ extracted from V_0; output recolored image frames $I_1', I_2', \ldots I'_k$; fixation data $F_1^l F_2^l, \ldots F_k^l$ extracted from D_0; alpha-blend image recoloring parameter α; eye-tracking metrics data storage arrays M and category of objects-of-interest C.

for k from 1 to K do:
  Perform image semantic segmentation on I_k;
    Δ Look up if F_k^l belongs to any segments of I_k;
  if TRUE then:
    Update M_k via calculation of the eye-tracking metrics;
    Update C_k via automatic categorization of fixated objects-of- interest;
  end
  Update I_k to I_k' via solving Eq.(1);
end

Output: $I_1', I_2', \ldots I'_k$; M; C and data representation ISeeColor including all the processed information.

scene video is required through eye-tracker's internal clock, since the infrared camera and front scene camera run with different frequencies. The code to produce fixation location in each video frame is written in python and can be downloaded from our research website.

Further, we denote α as the alpha-blend parameter for the image recoloring step. In addition, M is used to store calculated eye-tracking statistical metrics. For instance, metrics such as fixation duration, fixation rate, and total fixation number are measured and presented in our work. C is a vector to record the category of a fixated OOI for each frame.

B. IMAGE SEMANTIC SEGMENTATION USING DEEP LEARNING

As the authors were unable to find any public semantic segmentation datasets with fixation data, they were unable to train the dataset with eye fixations. As a work around, in this method, the authors have used state-of-the-art segmentation methods and applied fixation masks from the eye-tracker to get the object-of-interest.

The proposed ISeeColor is capable of automatically detecting, categorizing, and highlighting the OOI in a dynamic scene. ISeeColor architecture uses the correlation of fixation coordinates along with segments (objects) in a frame to identify the fixated OOI. Then the fixated OOI is overlaid by a transparent color mask with color intensity changing as fixation duration increases.

Image semantic segmentation describes the process of associating each pixel of an image with a class label, such as a car, truck, person, or dog. Deep convolutional neural networks based semantic segmentation clusters parts of the image together that belong to the same object. This technique has gained increasing attention from the research community and industry. To demonstrate the capabilities of ISeeColor, this system will use DeepLabv3 [45]–[49] and Context Encoding Network [50].

Chen *et al.* [49] developed DeepLabV3 which was designed originally for image classification, and it is remodeled to perform semantic segmentation. First, all the fully connected layers of the original network are transformed into convolutional layers, and then increasing feature resolution is increased through Atrous convolution. Up-sampling via bi-linear interpolation is applied to recover the original image resolution. This provides the input to a fully-connected Conditional Random Field [51] that improves the segmentation results.

The Intersection of Union (IoU) is the ratio between the area of overlap and the area of union between the ground truth and the predicted areas. The mean IoU (mIoU) is the average between the IoU of the segmented objects over all the images of the test dataset. mIoU is one of the most used commonly used metric in segmentation challenges. DeepLabv3 network produced an mIOU score of 85.7 in the 2012 PASCAL VOC challenge.

Zhang *et al.* [50] developed Context Encoding Network (EncNet) which captures global image information to achieve good scene segmentation. The model uses a ResNet to extract the feature maps which are then fed into a Context Encoding Module developed by Zhang *et al.* [52]. EncNet network produced an mIOU score of 85.9 in the 2012 PASCAL VOC challenge.

C. OOI COLOR TRANSFORM USING ALPHA-BLENDING

The image semantic segmentation algorithm will generate a group of segments, $\{S_1, S_2, \ldots, S_n\}$, where n is the number of possible OOI categories. For each video frame, a pixel-based linear operation is implemented as described in the equation below:

$$I' = \begin{cases} \alpha I + (1-\alpha)\, I_{color}; & I\,(x, y) \in \bigcup_{i=1}^{n} S_i \\ I; & otherwise \end{cases} \quad (1)$$

Here I' is the color blended destination image. I is the original eye-tracking video frame. I_{color} is the RGB color selected by fixation duration. $0 < \alpha < 1$ is used to control the transparency of overlaid color; in this work, $\alpha = 0.3$.

D. ISEECOLOR DATA VISUALIZATION AND ANALYSIS GUI

This article demonstrates the robust capabilities of ISeeColor through a user-friendly graphical user interface (GUI), where the OOI is recolored based on the fixation duration that can be directly displayed on a video stimulus. Users can instantly gain insights from the recorded visual stimuli. Currently, researchers will be able to choose between DeepLabV3 and EncNet for semantic segmentation. The authors will add more state-of-the-art algorithms in the future to provide a wider arsenal. The object categories can be customized based on researchers' needs.

The eye-tracker's analysis software provides the fixation location and fixation duration. Higher fixation durations are thought to indicate cognitive functions, such as active attentional deployment, while lower durations are considered to reflect the visual complexity of a scene [53]. ISeeColor also includes features to (1) observe histogram plots of the objects viewed as the video is played, (2) search for frames with an object, and (3) filter objects with high fixation duration.

The researcher can choose eye-tracking metrics that are relevant to the tasks and current research questions. As an illustrative example, this paper displays the features listed in Table 3 for ISeeColor GUI.

TABLE 3. Media statistics in ISeeColor.

Total time		A metric to record the task time
Total frames		Number of frames = time × frames /second
Total fixations		Provided by a commercial eye-tracker
Fixation rate	OOI	$\text{Fixation Rate}_{\text{objects}} = \frac{\#\text{ of fixations}_{\text{objects}}}{\text{Total time}}$
	No OOI	$\text{Fixation Rate}_{\text{No objects}} = \frac{\#\text{ of fixations}_{\text{No objects}}}{\text{Total time}}$
Number of fixations for OOI		Number of fixations the viewer spent on OOI

Furthermore, ISeeColor generates a histogram plot that is automatically updated as the frames are updated (in the forward or reverse direction). Moreover, the pertinent information regarding the existence of fixations in a video frame and its location are also displayed.

Figure 7 shows how much time an inexperienced and an experienced coder spent when they manually annotated two case study videos (video one is 1 minute 47 seconds, and video two is 3 minutes). The inexperienced coder was a research assistant who received a one-hour tutorial on using the proprietary annotation software from the eye-tracker's manufacturer (SMI BeGaze version 3.7, Sensomotoric Instruments, Inc.). The experienced coder was an eye-tracking researcher who has approximately six years of experience with mobile eye-tracking recordings and analysis.

FIGURE 6. Image Blending on segmented masks. The color is based on fixation counts.

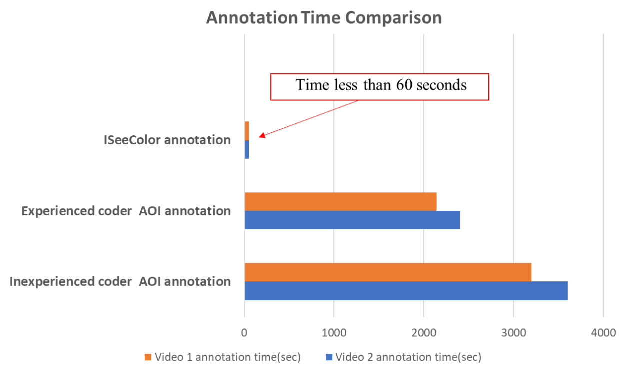

FIGURE 7. Comparison of annotation: human labor time between manually ROI annotation and ISeeColor. Notice, the ISeeColor annotation time is significantly minuscule compared to manual annotation methods.

Meanwhile, ISeeColor automated annotation and recoloring process did not require any human labor after feeding the data into the system, which takes less than 60 seconds.

Thus, the proposed ISeeColor integrates gaze direction information (where), enables automatic OOI labeling using image semantic segmentation (what), and facilitates visualization of data using fast-speed image recoloring based on the fixation duration (how long).

IV. USER-STUDY AND COMPARISON

We invited 51 eye-tracking researchers and non-researchers (including individuals from cognitive psychology, electrical engineering, management studies, computer science, industry professionals, and schoolteachers) to take part in a data visualization and analysis tool survey of ISeeColor. Each participant was given a short overview of eye-tracking, its challenges, and the current methods of annotation. Each participant was further shown a short demo of ISeeColor, that included an overview of all its functions and how it works.

Each question in the survey could be rated from 0 (extremely useless) – 10 (extremely useful). Table 4 shows the results of the survey. The high mean values of the results for all questions in the survey prove the effectiveness of the proposed system. Furthermore, the survey also asked for suggestions that could improve the visualization and the analysis process to solve the given tasks. Some of the feedback included interest in adding abilities to, (1) click an object-of-interest and to get the frames filtered from the video,

TABLE 4. Results of the ISeeColor user survey. Participants included researchers and non-researchers (professionals) from academia and industry. each participant was given an overview of the eye-tracking, its challenges, and a demo of ISeeColor.

Total number of participants	51		
Total number of participants who are familiar with eye-tracking technology	30		
Based on the ISeeColor demo, following aspects of the software were rated as follows:	**Mean**	**Std**	**Var**
The information is clear, concise, and informative to the intended audience	8.32	1.28	1.64
The purpose of the software is well-defined and clearly explained to the user	8.49	1.58	2.51
The organization of the software is clear, logical, and effective making it easy for the intended audience to understand	8.62	1.35	1.83
Computer capabilities such graphics, color, or sound are used for appropriate instructional reasons	8.64	1.26	1.6
Based on the ISeeColor demo, usefulness of following parameters or control options of the software were rated as follows:	**Mean**	**Std**	**Var**
Color-scale setting	8.51	1.58	2.51
Video control panel	8.79	1.29	1.67
Media Statistics	8.31	1.74	3.04
Histogram plot	8.15	2.16	4.68

(2) change the color scale to accommodate few cases of color blindness or color aphasia, and (3) allowing the user to adjust fixation count ranges to increase flexibility. The authors will include these valuable changes in future releases of the software.

Table 5 provides a brief comparison of popular eye-tracking data visualization and analysis approaches created for researchers to investigate eye movement information. Overall, while wearing head-mounted eye-trackers in dynamic contexts, ISeeColor can effectively and intuitively represent what, where, and how long users view objects or areas in a scene.

TABLE 5. A comparison of popular eye-tracking data visualization and analysis approaches.

Related work	Automatic OOI annotation	Fixation information presentation	
		Location	*Duration*
[30]		✓	✓
[31]		✓	✓
[22]		✓	✓
[32]			✓
ISeeColor	✓	✓	✓

V. EXTENDED APPLICATIONS FOR REAL-WORLD DYNAMIC CONTEXTS

ISeeColor has several potential applications for research and industry. The extraction of meaningful inferences from FoV scene videos and eye-tracking data during real-world locomotion remains a time-consuming and challenging process. However, these inferences can yield enormously helpful insights.

Consider researchers who wish to evaluate how well the cognitive theories that they developed in controlled laboratory settings apply in real-world dynamic contexts. Or consider the narrow problem where higher fixation durations could indicate (1) deeper cognitive processing of a stimulus, (2) greater difficulty in extracting information from the stimulus, or (3) that the stimulus is highly engaging. Researchers can use ISeeColor to detect and infer these scenarios in a matter of seconds.

In another instance, ISeeColor can be used to determine the extent to which someone looks at street signs, building, roads, or their smartphone map while walking around a city. This information can then be related to their subsequent memory for the environment.

Moreover, ISeeColor functionality could be applied to large-scale studies focused on understanding differences in viewing behavior between neurotypical populations and people with learning disorders or suffering from schizophrenia or

Alzheimer's. Consequently, ISeeColor has great potential for informing our understanding of human vision in naturalistic settings in an intuitive and user-friendly way.

In industry, ISeeColor can be used in product design to maximize profit, user satisfaction, and safety. For example, designers of semi-autonomous vehicles can examine the impact of system design choices on drivers' visual attention to critical OOI while driving. Another example includes market researchers who can determine what products catch shoppers' eyes during store visits to optimize product placement on shelves [53]. ISeeColor could open up new avenues for work in human-machine interaction and smart devices' user experience.

VI. CONCLUSION AND FUTURE WORK

This work proposed a novel visualization and analysis tool, ISeeColor, which enables large-scale wearable automatic eye-tracking data annotation and visualization. To the best of our knowledge, ISeeColor is the first eye-tracking data visualization and analysis system that can automatically visualize fixation events (duration and location) on a semantic interpretation of areas-of-interest in dynamic video recordings. ISeeColor includes implementation of state-of-the-art computer vision and deep learning techniques to automatically recognize and annotate fixated objects in FoV. ISeeColor utilizes color to represent fixation duration values. Using the ISeeColor GUI, OOI's are segmented and recolored with color intensity changing based on the fixation duration in an eye-tracking video recording. Finally, we have attempted to prove the effectiveness of the proposed method by conducting experiments with 51 participants from academia and industry including researchers and non-researchers.

ACKNOWLEDGMENT

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Combat Capabilities Development Command, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

[1] H. Ashraf, M. H. Sodergren, N. Merali, G. Mylonas, H. Singh, and A. Darzi, "Eye-tracking technology in medical education: A systematic review," *Med. Teacher*, vol. 40, no. 1, pp. 62–69, Jan. 2018.

[2] J. L. Orquin and K. Holmqvist, "Threats to the validity of eye-movement research in psychology," *Behav. Res. Methods*, vol. 50, no. 4, pp. 1645–1656, Aug. 2018.

[3] T. P. Keane, N. D. Cahill, J. A. Tarduno, R. A. Jacobs, and J. B. Pelz, "Computer vision enhances mobile eye-tracking to expose expert cognition in natural-scene visual-search tasks," in *Proc. 19th Hum. Vis. Electron. Imag.*, vol. 9014, Feb. 2014, p. 90140.

[4] K. M. Evans, R. A. Jacobs, J. A. Tarduno, and J. B. Pelz, "Collecting and analyzing eye tracking data in outdoor environments," *J. Eye Movement Res.*, vol. 5, no. 2, p. 6, 2012.

[5] M. Sodhi, B. Reimer, J. L. Cohen, E. Vastenburg, R. Kaars, and S. Kirschenbaum, "On-road driver eye movement tracking using head-mounted devices," in *Proc. Symp. Eye Tracking Res. Appl. (ETRA)*, 2002, pp. 61–68.

[6] H. Song, J. Lee, T. J. Kim, K. H. Lee, B. Kim, and J. Seo, "GazeDx: Interactive visual analytics framework for comparative gaze analysis with volumetric medical images," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 311–320, Jan. 2017.

[7] M. Vidal, J. Turner, A. Bulling, and H. Gellersen, "Wearable eye tracking for mental health monitoring," *Comput. Commun.*, vol. 35, no. 11, pp. 1306–1311, Jun. 2012.

[8] R. Netzel, M. Burch, and D. Weiskopf, "Comparative eye tracking study on node-link visualizations of trajectories," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2221–2230, Dec. 2014.

[9] S.-H. Kim, Z. Dong, H. Xian, B. Upatising, and J. S. Yi, "Does an eye tracker tell the truth about visualizations?: Findings while investigating visualizations for decision making," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2421–2430, Dec. 2012.

[10] S. Rajeev, Q. Wan, K. Yau, K. Panetta, and S. S. Agaian, "Augmented reality-based vision-aid indoor navigation system in GPS denied environment," in *Proc. Mobile Multimedia/Image Process., Secur., Appl.*, May 2019, Art. no. 109930.

[11] A. Al-Rahayfeh and M. Faezipour, "Eye tracking and head movement detection: A state-of-art survey," *IEEE J. Transl. Eng. Health Med.*, vol. 1, Nov. 2013, Art. no. 2100212.

[12] B.-S. Yoo and J.-H. Kim, "Evolutionary fuzzy integral-based gaze control with preference of human gaze," *IEEE Trans. Cognit. Develop. Syst.*, vol. 8, no. 3, pp. 186–200, Sep. 2016.

[13] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behav. Res. Methods, Instrum., Comput.*, vol. 34, no. 4, pp. 455–470, Nov. 2002.

[14] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "Scikit-image: Image processing in Python," *PeerJ*, vol. 2, p. e453, Jun. 2014.

[15] Q. Wan, K. Panetta, A. Kaszowska, H. Taylor, S. S. Agaian, and S. Rajeev, "Fixation oriented object segmentation using mobile eye tracker," in *Proc. Mobile Multimedia/Image Process., Secur., Appl.*, vol. 10668, May 2018, Art. no. 106680.

[16] G. Andrienko, N. Andrienko, M. Burch, and D. Weiskopf, "Visual analytics methodology for eye movement studies," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2889–2898, Dec. 2012.

[17] P. Isokoski, J. Kangas, and P. Majaranta, "Useful approaches to exploratory analysis of gaze data: Enhanced heatmaps, cluster maps, and transition maps," in *Proc. ACM Symp. Eye Tracking Res. Appl.*, 2018, Art. no. 68.

[18] J. Steil, M. X. Huang, and A. Bulling, "Fixation detection for head-mounted eye tracking based on visual similarity of gaze targets," in *Proc. ACM Symp. Eye Tracking Res. Appl. (ETRA)*, 2018, p. 23.

[19] H. Y. Tsang, M. Tory, and C. Swindells, "eSeeTrack—Visualizing sequential fixation patterns," *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 6, pp. 953–962, Nov./Dec. 2010.

[20] K. Kurzhals, F. Heimerl, and D. Weiskopf, "ISeeCube: Visual analysis of gaze data for video," in *Proc. Symp. Eye Tracking Res. Appl.*, 2014, pp. 43–50.

[21] R. Netzel, M. Burch, and D. Weiskopf, "Interactive scanpath-oriented annotation of fixations," in *Proc. 9th Biennial ACM Symp. Eye Tracking Res. Appl. (ETRA)*, 2016, pp. 183–187.

[22] K. Kurzhals, M. Hlawatsch, C. Seeger, and D. Weiskopf, "Visual analytics for mobile eye tracking," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 301–310, Jan. 2017.

[23] P. Yin, H. Qiao, W. Wu, L. Qi, Y. Li, S. Zhong, and B. Zhang, "A novel biologically inspired visual cognition model: Automatic extraction of semantics, formation of integrated concepts, and reselection features for ambiguity," *IEEE Trans. Cognit. Develop. Syst.*, vol. 10, no. 2, pp. 420–431, Jun. 2018.

[24] D. F. Pontillo, T. B. Kinsman, and J. B. Pelz, "SemantiCode: Using content similarity and database-driven matching to code wearable eyetracker gaze data," in *Proc. Symp. Eye-Tracking Res. Appl.*, 2010, pp. 267–270.

[25] T. Nakamura and T. Nagai, "Ensemble-of-concept models for unsupervised formation of multiple categories," *IEEE Trans. Cognit. Develop. Syst.*, vol. 10, no. 4, pp. 1043–1057, Dec. 2018.

[26] E. Kowler, "Eye movements: The past 25 years," *Vis. Res.*, vol. 51, no. 13, pp. 1457–1483, 2011.

[27] L. R. Squire, N. Dronkers, and J. Baldo, *Encyclopedia of Neuroscience*. Amsterdam, The Netherlands: Elsevier, 2009.

[28] A. Poole and L. J. Ball, "Eye tracking in HCI and usability research," in *Encyclopedia of Human Computer Interaction*. Hershey, PA, USA: IGI Global, 2006, pp. 211–219.
[29] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: Methods and constructs," *Int. J. Ind. Ergonom.*, vol. 24, no. 6, pp. 631–645, Oct. 1999.
[30] K. Kurzhals and D. Weiskopf, "Space-time visual analytics of eye-tracking data for dynamic stimuli," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2129–2138, Dec. 2013.
[31] M. Raschke. *Eye Tracking—The Next Big Thing for Automatic Driving*. Accessed: Jun. 10, 2019. [Online]. Available: https://www.visual-computing.org/2016/08/16/eye-tracking-the-next-big-thing-forautomatic-driving/
[32] R. Jianu and S. S. Alam, "A data model and task space for data of interest (DOI) eye-tracking analyses," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 3, pp. 1232–1245, Mar. 2018.
[33] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl, "State-of-the-art of visualization for eye tracking data," in *Proc. EuroVis*, 2014, pp. 1–20.
[34] J. H. Goldberg and J. I. Helfman, "Visual scanpath representation," in *Proc. Symp. Eye-Tracking Res. Appl. (ETRA)*, 2010, pp. 203–210.
[35] D. S. Wooding, "Fixation maps: Quantifying eye-movement traces," in *Proc. Symp. Eye Tracking Res. Appl. (ETRA)*, 2002, pp. 31–36.
[36] R. Netzel, M. Hlawatsch, M. Burch, S. Balakrishnan, H. Schmauder, and D. Weiskopf, "An evaluation of visual search support in maps," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 421–430, Jan. 2017.
[37] A. T. Duchowski, J. Driver, S. Jolaoso, W. Tan, B. N. Ramey, and A. Robbins, "Scanpath comparison revisited," in *Proc. Symp. Eye-Tracking Res. Appl. (ETRA)*, 2010, pp. 219–226.
[38] T. Opach and A. Nossum, "Evaluating the usability of cartographic animations with eye-movement analysis," in *Proc. 25th Int. Cartographic Conf.*, Jul. 2011, p. 11.
[39] A. Çöltekin, S. I. Fabrikant, and M. Lacayo, "Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings," *Int. J. Geographical Inf. Sci.*, vol. 24, no. 10, pp. 1559–1575, Oct. 2010.
[40] K. Kurzhals, M. Hlawatsch, F. Heimerl, M. Burch, T. Ertl, and D. Weiskopf, "Gaze stripes: Image-based visualization of eye tracking data," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 1, pp. 1005–1014, Jan. 2016.
[41] T. Blascheck, M. John, K. Kurzhals, S. Koch, and T. Ertl, "VA^2: A visual analytics approach for evaluating visual analytics applications," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 1, pp. 61–70, Jan. 2016.
[42] P. K. Muthumanickam, K. Vrotsou, A. Nordman, J. Johansson, and M. Cooper, "Identification of temporally varying areas of interest in long-duration eye-tracking data sets," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 87–97, Jan. 2019.
[43] S. S. Alam and R. Jianu, "Analyzing eye-tracking information in visualization and data space: From where on the screen to what on the screen," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 5, pp. 1492–1505, May 2017.
[44] G. Intelligence. (2019). *Solutions for Behavioral Studies, Psychology and Multimodal Research*. [Online]. Available: https://gazeintelligence.com/smi-product-manual
[45] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," 2018, *arXiv:1802.02611*. [Online]. Available: http://arxiv.org/abs/1802.02611
[46] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," 2016, *arXiv:1610.02357*. [Online]. Available: http://arxiv.org/abs/1610.02357
[47] J. Dai, "Deformable convolutional networks," *CoRR*, vol. 06211, no. 2, p. 3, 2017.
[48] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.
[49] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*. [Online]. Available: http://arxiv.org/abs/1706.05587
[50] H. Zhang, K. Dana, J. Shi, Z. Zhang, X. Wang, A. Tyagi, and A. Agrawal, "Context encoding for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7151–7160.
[51] C. Sutton, "An introduction to conditional random fields," *Found. Trends Mach. Learn.*, vol. 4, no. 4, pp. 267–373, 2012.
[52] H. Zhang, J. Xue, and K. Dana, "Deep TEN: Texture encoding network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 708–717.
[53] A. T. Duchowski, "Eye tracking methodology," in *Theory and Practice*, vol. 328. Cham, Switzerland: Springer, 2007.

KAREN PANETTA (Fellow, IEEE) received the B.S. degree in computer engineering from Boston University, Boston, MA, USA, and the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Boston. She is currently the Dean of the Graduate Engineering Education, a Professor with the Department of Electrical and Computer Engineering, an Adjunct Professor of computer science with Tufts University, Medford, MA, and the Director of the Dr. Panetta's Vision and Sensing System Laboratory. She is the President-Elect 2019 of IEEE-HKN. Her researches focus on developing efficient algorithms for simulation, modeling, signal, and image processing for biomedical and security applications.

QIANWEN WAN (Student Member, IEEE) received the B.S. degree in information engineering and the B.S. degree in law from the Wuhan University of Technology, Wuhan, China, and the M.S. and Ph.D. degrees in electrical and computer engineering from Tufts University, Medford, MA, USA.

SRIJITH RAJEEV (Student Member, IEEE) received the B.E. degree in electronics and communication engineering from Visvesvaraya Technological University, India, in 2014, and the M.S. degree in electrical and computer engineering from The University of Texas at San Antonio, USA, in 2016. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Tufts University, USA. His current research interests include artificial intelligence, eye-tracking, signal/image processing, machine learning, 3-D sensors and modeling, digital forensic, and biomedical applications.

ALEKSANDRA KASZOWSKA received the B.A. degree in psychology from Clark University, Worcester, MA, USA, and the M.S. degree in experimental psychology from Tufts University, Medford, MA, where she is currently pursuing the Ph.D. degree in cognitive science. Her current research interests include image and video processing, eye tracking, problem solving, and spatial cognition.

AARON L. GARDONY received the B.A. degree in psychology and the joint Ph.D. degree in psychology and cognitive science from Tufts University. He is currently a Cognitive Scientist with the Cognitive Science and Applications Team, U.S. Army Combat Capabilities Development Command Soldier Center (CCDC SC), and a Visiting Scholar with the Center for Applied Brain and Cognitive Sciences (CABCS). His current research interests include spatial cognition, augmented/virtual reality, and human–computer interaction.

KEVIN NARANJO is currently pursuing the bachelor's degree in electrical engineering and computer science with Tufts University, Medford, MA, USA. He is also working as a Research Assistant with the Dr. Panetta's Vision and Sensing Laboratory. His principle responsibilities involve developing the software architecture to analyse and visualize large-scale eye-tracking data. He has also worked on other projects relating to facial recognition and deep learning.

HOLLY A. TAYLOR received the B.A. degree in mathematics from Dartmouth College, Hanover, NH, USA, and the Ph.D. degree in cognitive psychology from Stanford University, Stanford, CA, USA. She is currently a Full Professor of psychology and an Adjunct Professor of mechanical engineering, and the Co-Director of the Center for Applied Brain and Cognitive Science, Tufts University.

SOS AGAIAN (Fellow, IEEE) is currently a Distinguished Professor with The City University of New York/CSI. His research interests include computational vision and machine learning, large scale data analytic analytics, multimodal data fusion, biologically-inspired signal/image processing modeling, multimodal biometric and digital forensics, 3-D imaging sensors, information processing, and security, and biomedical and health informatics. He is an Editorial Board Member for the *Journal of Pattern Recognition and Image Analysis* and an Associate Editor for several journals, including the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS, the *Journal of Electrical and Computer Engineering* (Hindawi Publishing Corporation), the *International Journal of Digital Multimedia Broadcasting* (Hindawi Publishing Corporation), and *Journal of Electronic Imaging* (SPIE and IS&T).