Deep Learning-based Eye Gaze Estimation for Military Aviation

L R D Murthy
I3D Lab, CPDM,
Indian Institute of Science,
Bangalore,
lrdmurthy@iisc.ac.in

Pradipta Biswas
I3D Lab, CPDM,
Indian Institute of Science,
Bangalore,
lrdmurthy@iisc.ac.in

Abstract—**Eye gaze estimation and cognitive load estimation of the pilot garnered great attention in the aviation domain due to the numerous possible applications. Earlier works proposed to use eye gaze tracking to interact with multi-function displays (MFDs) and head-up display (HUD) in place of traditional interaction devices. Further, researchers also investigated the accuracy of commercially available gaze trackers during in-flight conditions by conducting studies under various actual flying scenarios like varying g-conditions and different maneuvers. In this paper, we first studied the functioning of a wearable eye gaze tracker using two one-hour long flights. Pilots undertook various challenging maneuvers during the flight. We analyzed the gaze tracking data recorded using the gaze tracker and observed that the ~42% and ~31% of flight duration resulted in loss of gaze data in flight1 and flight 2 respectively. Further, we analyzed unsynchronized raw data and observed that both flights recorded error-prone gaze samples for ~51% of the flight duration. We hypothesized and verified that this loss of data is caused due to the higher levels of illumination on eyes and limited field of view provided by the gaze tracker in the vertical direction. The data from both flights supported our hypothesis and it was evident that the current field of view offered by the eye tracking glasses is not sufficient for the military aviation. We addressed the first limitation by using Machine learning approach. We built an end-to-end gaze estimation system which takes IR-eye images recorded using wearable eye tracking glasses to predict the gaze point. We sampled 10K images with proper ground truth gaze points. Our dataset contained wide variation in illumination and pupil dilation. We observed that the proposed approach using a convolutional neural network resulted in low gaze estimation errors and consistent gaze predictions.**

TABLE OF CONTENTS

978-1-6654-3760-8/22/$31.00 ©2022 IEEE

1. INTRODUCTION

Eye gaze estimation research has witnessed numerous applications in recent times. In addition to the applications like assistive technology [1] , eye scan path analysis [2] and Usability research [3], eye gaze tracking is being adopted to wider domains like automotive to estimate cognitive load of the driver [4] and to operate infotainment screens inside the cars [5]. We observe that eye tracking technology is pervading the aviation domain with similar applications like cognitive load estimation of the pilot and for operating multi-functional displays (MFDs) and Head-up Displays (HUDs) of the aircrafts [6].

Even though earlier research demonstrated the utility of this technology in aviation domain, we observed little focus on performing the feasibility study inside an actual flying environment. Since the actual flying environments involve variations in terms of external illumination, altitude and pilot's cognitive state which in turn affects pupil dilation, it is still unclear about the reliability and consistency of the commercially available off the shelf (COTS) eye trackers in such conditions. Further, infrared-based commercial wearable eye trackers did not report accuracy in outdoor conditions and beyond 3000 lux [7] . In this paper, we performed the failure mode analysis of a COTS eye tracker in an actual flying environment using two flight recordings. We recorded data in two fighter planes with two different pilots and we observed that COTS eye tracker could not record eye gaze for ~42% and ~31% of flight duration respectively. We investigated this loss of gaze data occurred in two flights. We proposed two hypotheses to explain the loss of gaze data and we corroborated our hypotheses by analyzing the recorded gaze data. Further, we also proposed an end-to-end gaze estimation system using machine learning. We observed that the proposed machine-learning based gaze estimation system could provide gaze estimates more consistently than COTS eye trackers which rely on feature-based methods.

In the next section, we presented our literature review related to eye gaze tracker development and evaluation. In section 3, we presented our failure mode analysis of a COTS eye tracker. In section 4, we presented the design and development of IR-based gaze estimation system using the data recorded in actual flying conditions. In section 5, we

2022 IEEE Aerospace Conference (AERO) | 978-1-6654-3760-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/AERO53065.2022.9843506

discussed limitations and future work followed by conclusion in section 6.

2. RELATED WORK

Infrared-based approach was widely used for eye gaze estimation compared to appearance-based approach due to its well-established theory using pupil centers and corneal reflections. These systems used feature-based approach for gaze estimation by extracting features using multiple IR illuminators and cameras. They also employed a calibration procedure to obtain user-specific parameters which contributes to reportedly higher accuracy of this approach.

Numerous commercially available eye trackers utilized this feature-based approach with IR illuminators to obtain high accuracies. Technical specifications of a screen-based eye trackers [8] reported <18 mm error across >95% of population under real world usage conditions. We observed similar performance was not reported for wearable eye trackers which provided unconstrained gaze estimation irrespective of the environment. Technical specification document of Tobii Pro glasses 2 [7] reported 0.62° error for an evaluation conducted in-house using 20 participants for gaze angles ≤15° and 3.05° ± 1.13° for gaze angles larger than 15°. Further, even though these eye trackers offered unconstrained gaze estimation, makers revealed that the optimal operating conditions are 300 lux with gaze angles <15° and target lying at a distance less 1.5m from the user. They reported an increased gaze estimation error of 0.79° at 3000 lux compared to 0.62° at 300 lux and no evaluation was performed beyond 3000 lux.

In addition to Tobii Pro Glasses 2, we also reviewed the existing evaluation on Pupil Invisible [9], another commercially available wearable eye tracking glasses. The latter employed lesser number of cameras and IR illuminators per eye. The evaluation performed on Pupil Invisible is larger and more inclusive than the one performed using Tobii [7]. Pupil invisible evaluation consisted of 367 unique participants with gaze angles ranged from around ±45° in both pitch and yaw directions. The data was recorded in a depth range of 30-350 cm. Authors reported that the data recorded included one indoor and one outdoor environment for each participant and the outdoor environment ranging from cloudy to sunny. Authors reportedly observed no significant difference between the gaze estimation errors obtained indoor and outdoor conditions, but they reported higher error per subject (5.5° ± 2°) than Tobii Pro glasses 2. Further, even though they reported to include outdoor conditions as sunny, range of illumination intensity was not reported.

Numerous efforts were also made on research front towards developing IR-based gaze estimation systems. Guestrin and Eizenman [10] reported to have obtained a PoG (Point of Gaze) accuracy of 0.9° in a desktop setting based on an evaluation on 4 subjects. Further, Brousseau et al [11] proposed a system for gaze estimation for mobile devices compensating for the relative roll between the system and subject's eyes. They evaluated their system on 4 subjects and reported around 1° of gaze estimation error. Even though these results were promising, it was unclear how these systems will perform in outdoor conditions with high illumination, as in the case of actual flying conditions.

Summarizing our literature review, we found little evidence of evaluating wearable eye gaze trackers in bright outdoor conditions higher than 3000 lux. Further, existing IR-based gaze estimation systems were predominantly feature-based approaches. Since environment do have an impact on the data and in-turn on the machine learning model's performance, we observed few machine learning approaches trained on the data collected from actual flying conditions.

3. FAILURE ANALYSIS OF COTS EYE TRACKER

Our earlier evaluation [12] of eye gaze tracking technology in military aviation involved both transport and fighter aircrafts. We evaluated the accuracy of a screen-based COTS eye tracker in actual flying conditions for a pointing and selection task. We conducted our evaluation on ground and in-air conditions inside a transport aircraft. We compared adaptive selection strategy, which utilized nearest-neighborhood algorithm to activate target nearest to gaze location with non-adaptive condition. We noted that the pilots preferred adaptive selection strategy compared to non-adaptive condition. During the study of transport aircraft, we observed that the screen-based eye tracker did not fail to perform gaze detection. We also conducted a study on fighter aircraft using wearable eye gaze tracker. We observed significant difference in the accuracy of eye gaze tracker across different 'g' conditions. For more details, we encourage reader to refer to [12].

During our evaluation of accuracy of eye gaze trackers in a fighter aircraft scenario, we observed that gaze points were not recorded for a significant flying time. We performed a detailed analysis to investigate and understand the scenarios in which this loss of gaze data can happen. We proposed two hypotheses to explain the loss of data while using the Tobii pro glasses 2 wearable eye tracker and we corroborated them using the gaze points data we recorded during the flight.

To this end, we recorded data from two flights using the COTS eye tracker (Tobii Pro Glasses 2) which used InfraRed (IR) illumination-based eye gaze estimation principles. We chose Tobii Pro Glasses 2 over Pupil Invisible due to its reportedly higher accuracy. The duration of the first flight was 55 minutes 58 seconds (Flight 1) and another flight's duration was 56 minutes (Flight 2). We furnished the flight profiles in Table 1. The eye tracker contained a front-facing scene camera which recorded the first person view of the pilot. It also contained four eye-cameras, two cameras per each eye, to record the eye movements. The eye tracker estimated gaze points at a frequency of 100 Hz. The frame rate of scene camera was 25.01 frames/second at 1920 x 1080 resolution and that of each eye camera was around 50 frames/second with a resolution of 240x240. Each gaze point was recorded with a dedicated identifier, called "gidx". We initially used Tobii Pro Lab tool to analyze the recorded gaze samples and observed that both flight recordings contain gaze samples only for around 50% of the duration. We investigated the reason behind this loss of data samples during the flight using the raw data provided by manufacturer in json format and by correlating the raw data with the eye images.

Table 1. Flight Profiles

S. No	Objective	Profile
Flight #1	Maneuvering flight with head mounted eye tracker on Pilot in Command	Take-off – climb – level flight to Local Flying Area – Constant G (3G and 5G) level turns both sides each – Vertical loop – Barrel Roll – Air to Ground dive attack training missions – Descent – Instrument Landing System (ILS) Approach and landing
Flight #2	Non - Maneuvering flight with head mounted eye tracker on Pilot in Command	Take-off – climb – level flight to Local Flying Area – Straight and Level cruise with gentle level turns – Descent – ILS Approach and landing

Data Pre-Processing

The raw data obtained in json format contained various other information recorded during the flight like gyroscope and accelerometer data. We discarded such irrelevant information and retained the data points required for our investigation of the failure modes of eye tracker.

At first, we synchronized the raw data stream and the eye camera stream in time scale since eye camera stream starts off with an offset from raw data. This was achieved using the Position Time Stamps (PTS) provided in both data streams. We also observed that the different frequencies of these two streams posed a challenge for data synchronization. Hence, we considered average of all gaze points recorded during the time window between two successive scene camera frames and assigned the average gaze point to the latter frame. Thus, the latter frame and these data points together formed one pair of synchronized data point. Each time windowed raw data may contain multiple gaze points. Every gaze point with its "gidx" contains a status code, 's' which indicates the error associated to that datapoint, if any. The status value 0 indicates no error and any non-zero value of s indicates an error associated with the data point. The gaze points provided by manufacturer were in normalized values; hence the minimum gaze point was [0.0, 0.0] and the maximum was [1.0,1.0].

We split the synchronized data points into two categories. The first category *category1* contained eye stream frames whose corresponding gaze points had zero status code. The second category *category2* contained those eye frames with all corresponding gaze points with non-zero status codes. There were frames whose data points had only a subset of gaze points with zero status code. We did not consider those frames in our analysis as it brought uncertainty on eye image tagging. In other words, *category1* contained camera frames with proper gaze points and *category2* contained frames with no gaze points.

For Flight 1, we observed that out of 167,647 frames, only 57,111 frames come under *category1* and 69,732 frames come under *category2*. For Flight 2, we observed that out of 167,567 frames, only 81,911 frames fall under *category1* and 51,402 frames fall under *category2*.

Summarizing, 41.6% of the frames did not have any gaze points recorded during Flight 1 and for Flight 2, this stood at 30.7%. Further, if we considered unsynchronized raw data alone, both flights recorded more than 51% of the gaze samples were error-prone.

We visually inspected these flight recordings and we hypothesized two reasons for this data loss.

1. *Higher levels of illumination on eyes may affect the eye tracker resulting in no gaze estimation.*
2. *Limited field of view (FoV), especially in the vertical direction, renders the eye tracker with no gaze estimates when user looks beyond the tracking range.*

Effect of Illumination

We validated our hypothesis 1 using the eye images in the above mentioned two categories. We converted all eye images into grayscale and computed average of all the pixel values for each image present in both *category1* and *category2*. Figure 1 represents the histogram of image intensities for *category1* and *category2* for Flight #1. Figure 2 represented the same for Flight 2. Figure 1a indicates that 93% of the images under *category1* had an average pixel value less than 131. But *category2* contained 42% with average intensity higher than 131. In other words, only 7% of the images with intensity higher than 131 had proper gaze points and over 42% of images with no gaze points had intensity higher than 131.

A similar phenomenon could be observed in Flight 2, shown in Figure 2. Around 42% in *category2* had higher intensity than 150, while *category1* contained 94% of the images with intensity less than 150. In this case as well, only 6% of the images with intensity higher than 150 had proper gaze points and over 42% of images with no gaze points had intensity higher than 150. This indicated that images with higher illumination, precisely above 131 in Flight 1 and above 150 in flight 2 had low probability to obtain accurate gaze estimates.

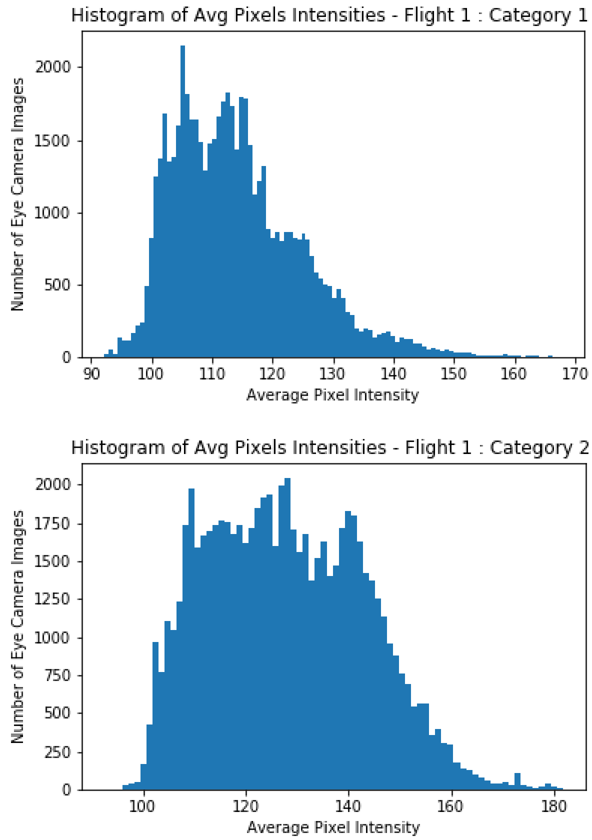

Figure 1. Histogram of image intensities for Flight #1

1a. *Category1*(Top) 1b. *Category2* (Bottom)

While this evidence supported our hypotheses 1 partially, we observed that there was overlap in the histograms plotted for both categories in Figure 1 and Figure 2. Hence, we could not identify a clear mean image intensity threshold to identify all the failure modes of eye gaze estimation.

Effect of Tracking Field of View

The illumination factor did not explain 58% of the images without proper gaze points in both flights data. We further investigated such data points in *category2* which had lower image intensities than above mentioned thresholds for each flight. During our visual inspection of first person video recorded using eye tracker, we observed that the pilot looked down for various activities like looking at the information displayed in the Multi-functional displays (MFDs). During such scenarios, we observed that gaze points were not recorded.

Extending our second hypotheses, we assumed that the pilot must had looked at a position closer to the extreme tracking positions (beyond which eye tracker could not track), just before or after the eye tracker failed to provide gaze estimates. Since we observed that the gaze estimates were lost for a sequence of eye image frames, we clustered the datapoints in *category2* based on their "gidx"s. If a sequence of datapoints under *category2* were having successive "gidx"s, then all those points were considered as a single cluster. Hence, each cluster can contain one datapoint or several datapoints.

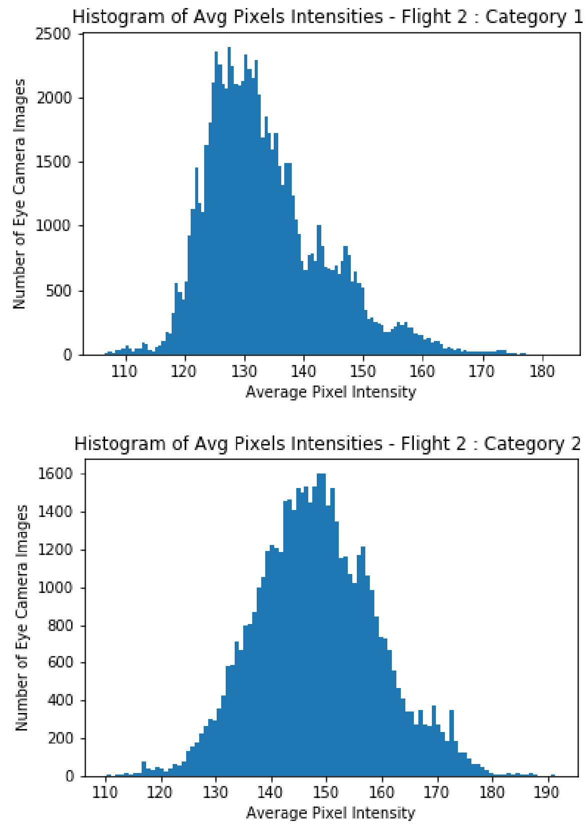

Figure 2. Histogram of image intensities for Flight #2

2a. *Category1*(Top) 2b. *Category2* (Bottom)

We analyzed three preceding or subsequent datapoints adjacent to each cluster, which we refer to as boundary datapoints. We looked for boundary datapoints with gaze values greater than 0.8 and less than 0.2 (in both x and y). If any of the boundary datapoints satisfy above criterion, then we may infer that the loss of gaze points was due to the pilot looking beyond the tracking range of the eye tracker.

For flight 1, we obtained 12178 clusters for 69,732 datapoints. For these clusters, 11,865 (97.43%) clusters had boundary points that satisfy the above criterion. To understand image intensities for these datapoints, we obtained the image intensities for boundary datapoints that satisfy above criterion. We observed that these image intensities lie in the range of (96,145). This was clearly in the overlap range identified between Figure 1a and Figure 1b.

Similarly, for flight 2, we obtained 8646 clusters for 51,402 datapoints. For these clusters, 8408 (97.24%) clusters had boundary points that satisfy the above criterion. Interestingly here as well, we observed that the image intensities for the above points lie in the range of (117,164), which was the range of overlap identified in Figure 2a and Figure 2b.

Thus, we infer that, the COTS eye gaze tracker could not identify beyond certain illumination level or if the user is looking beyond its tracking range. We note that the pilot is performing his assigned tasks during the flight and maintained his natural behavior. This indicates that the tracking range offered by the eye tracker used in this study falls short for military aviation environments.

Hence, using our two hypotheses, we explained the failure modes of an eye gaze tracker in aviation environment. We further add that, while commercial off-the shelf eye trackers can be useful in real aviation environments, researchers and practitioners should keep in mind about both the horizontal and vertical tracking range of the eye tracker and its robustness to external illumination as there is a high chance that the illumination varies rapidly at high altitudes and in high speed maneuvers.

4. Eye Gaze Estimation in High Illumination Environments

Deep Learning Model Design and Development

In this section, we elaborated the process of designing an end-to-end gaze estimation system using machine learning without relying on any handcrafted features.

We considered data obtained from both flights and carried out within person evaluation on each pilot. In order to avoid redundancy in learning, we sampled 10K images from each flight's *category1* as described in previous section. We ensured that our sampled data achieved similar data distribution as that of the entire data from respective flight recordings. In figures 3 and 4, we illustrated the sampled gaze direction distribution of both recordings. Since pilots had displayed their natural gaze behavior, the gaze directions were less spread across X-direction. We also noted that the gaze direction in flight #2 is wider than flight #1. We separated the sampled 10K data points from each flight into 3 splits, 6.4K, 1.6K and 2K for training, validation and testing respectively. Our dataset included wide variation of illumination and pupil dilation from our both flight recordings. We noted that the pupil dilation values in both flight recordings were in the range of 1.5 mm and 5.45 mm.

We performed both within person and cross-person evaluations on the two flights data.

We considered two factors while designing our model architecture. We considered both latency and accuracy as the design parameters. We used two convolutional layers with a kernel size of 3x3 and feature maps of 32 and 64 respectively. We also used three fully connected dense layers with 32, 16 and 4 neurons respectively. Further, we added a batch normalization layer after every layer in our model. We used an input image size of 240x960 which we obtained from Tobii eye camera recordings and our output layer consisted of two neurons representing gaze estimates in x and y

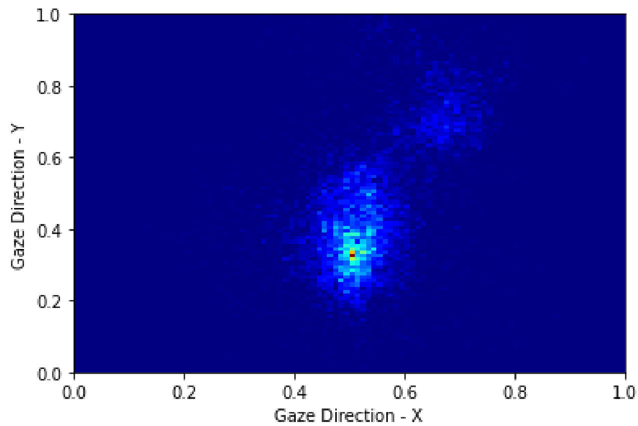

Figure 3. Histogram of Gaze Directions for Flight #1

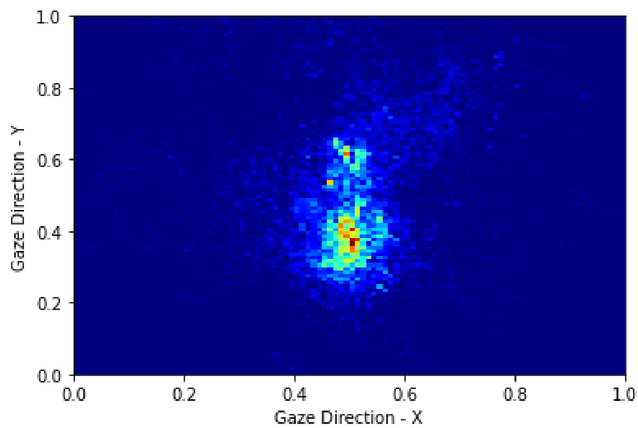

Figure 4. Histogram of Gaze Directions for Flight #2

direction. We obtained normalized gaze direction labels using Tobii with x and y gaze directions ranging from 0 to 1. We used mean square error as the loss function and Adam for the optimization. We used a batch size of 32 and we trained the model for 50 epochs. We monitored the validation loss parameter to save the best performing model.

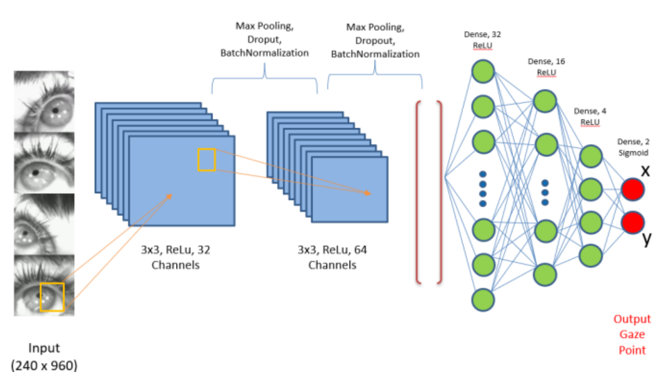

Figure 5. Network Architecture for Gaze Estimation

Table 2. Within Person Validation Results

#Flight	Test Error
1	4E-3
2	8E-3

Within Person Experiments

We presented the results of within person validation in Table 2. We observed that the test error obtained was in the order of E-3 for both participants. We achieved this performance using 6.4K training samples, which will be equivalent to 2.18 minutes of recording using Tobii's eye cameras (50 fps). Further, since the predicted gaze points were mapped to the scene camera of resolution 1920x1080, we transformed the normalized gaze predictions to pixels. We observed a mean pixel error of 87 and 127 for flights #1 and #2 respectively.

Cross-person Validation

We also conducted cross-person evaluations to understand the generalizability of the proposed approach for unseen scenarios and persons.

Table 3. Cross-Person Validation Results

#Flight	Test Error
1	19E-3
2	23E-3

For each flight data, we utilized the model trained for with-in person evaluation with best validation loss and we tested on the other flight data. We presented our cross-person validation results in Table 3.

We observed that the cross-person evaluation resulted in higher errors compared to within person evaluation. We posit that there were couple of factors behind the high cross-person errors. The first reason was the differences in the nose cap while wearing the glasses which alters the appearance of eye images which might affect the model. The second being the person-specific factors like inter-pupillary distance (IPD) and offset between visual and optical axis for each person. Further, domain gap between training and testing datasets also might result in higher validation error. We illustrated two images from two recordings while pilot is looking at the same point in figure 6. We observed offset in eye images from both recordings, specifically in right eye images. We intend to handle such appearance variations using larger dataset and data augmentation techniques in our future work.

Consistent Gaze Predictions

We aimed to address the inconsistency in gaze estimates obtained from COTS eye tracker using our proposed machine learning approach. Since we used eye gaze labels obtained from Tobii for training our model, we did not have ground

Figure 6. Sample Images from Two flight recordings

truth gaze labels for the scenarios where Tobii failed to detect eye gaze. As we explained in Section 2, we named such images as *category2*. We verified whether the gaze predictions obtained for these images were consistent with prior and post gaze estimates provided by Tobii.

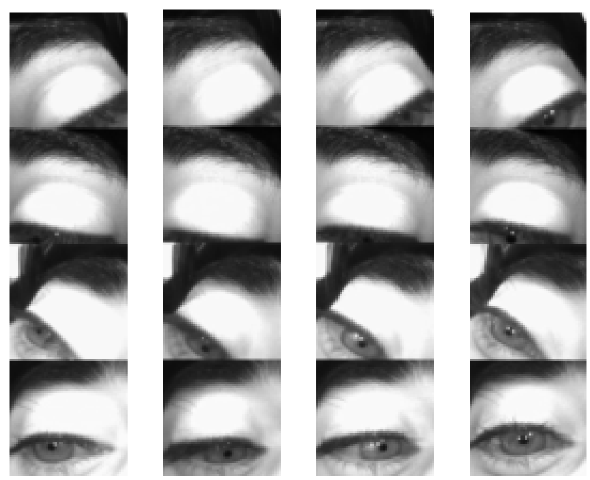

Figure 7. Sample Images from Time Sequence with No gaze estimates from COTS eye tracker

We obtained time sequences from Flight #1 recording where the loss of gaze data was longer than 10 seconds. We found 11 such instances and we considered the first sequence to illustrate the ability of the proposed approach to produce consistent gaze estimates. We obtained 391 eye images during the chosen time sequence and we predicted gaze points for these images. We illustrated a few sample images from this sequence in figure 7. It can be noted that left eye images in this sequence were partially occluded while right eye images indicated that the pilot was looking towards bottom half of the gaze plane.

In figure 8, we presented the gaze predictions obtained from 391 images for which COTS eye tracker failed to detect gaze. Our analysis for flight #1 in Section 2 identified 131 as the illumination threshold and the selected sequence of images displayed mean illumination higher than 137 as illustrated in figure 9.

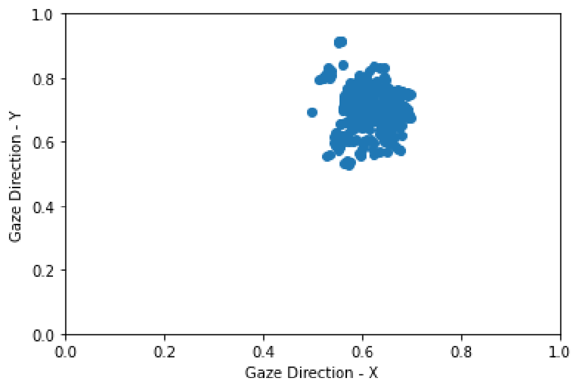

Figure 8. Gaze Predictions for Sequence of Images with no ground truth gaze labels

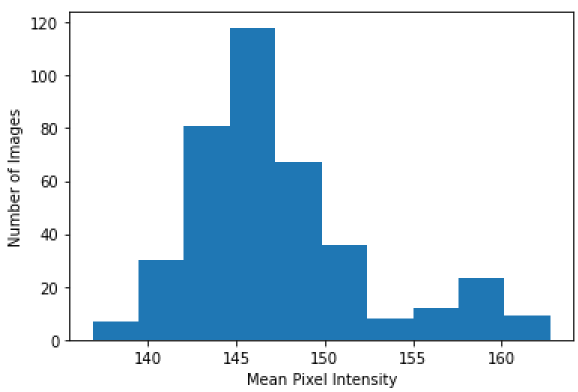

Figure 9. Histogram of Mean Pixel Intensity for Sequence of Images with no ground truth gaze labels

The prior and post gaze estimates for the sequence under consideration were [0.515, 0.567] and [0.51, 0.34] respectively. It can be observed that the predicted gaze points lie in the range of [0.49, 0.69] in X – direction and [0.52, 0.91] in Y – direction. Reiterating the two proposed hypotheses in Section 2 and inferring from the eye images in the selected sequence, we posit that pilot might also be looking at vertical extreme of tracking range which resulted in higher values of Y-direction of predicted gaze estimates. Hence, it was evident from the prior and post gaze estimates, eye camera images and the predicted gaze points, that the proposed method was consistent and can generalize well even for a high illumination setting where COTS eye tracker fail to detect eye gaze.

5. DISCUSSION

We presented our failure mode analysis of COTS wearable eye tracker. We recorded eye gaze tracking data using fighter aircrafts in actual flying conditions with two pilots. Our analysis indicated that the bright external illumination resulted in gaze detection failure. COTS eye tracker failed to detect gaze points for ~42% and ~31% of the flight duration. The technical specification of the COTS eye tracker mentioned that 5-10% of the data during a recording would be lost due to eye blinks. Hence, even after accounting for blinks during the recording, we observed significant data loss due to the failure of COTS eye tracker in actual flying conditions. Even though we verified our hypotheses only on two flight recordings, each of our recordings were one hour long. We intend to extend our analysis on more flight recordings as they become available. Feature-based approaches used by COTS eye trackers rely on the effectiveness of feature extraction algorithms. These extracted features include pupil centers, corneal reflections of IR illuminators from the eye images. Such systems tend to fail in high illumination conditions as the feature extraction systems could not distinguish between the corneal reflections caused by external IR illumination and inherent IR-illuminators of the system. Our second hypothesis highlight limited vertical field of view of the existing gaze trackers. We also presented the technical reports of the COTS eye tracker where higher gaze estimation errors were reported for higher gaze angles. We would like to highlight that the applications in aviation like cognitive load estimation of the pilot and operating secondary displays require higher vertical field of view for actual deployment.

We presented a machine learning based gaze estimation system which did not utilize any handcrafted features. One of the limitations of our architecture was the memory footprint. We intend to reduce it in our future work by experimenting with lower input image size, data augmentation and advanced deep learning techniques. Further, we evaluated our model only on two pilots' data as collecting such data itself is a challenge. We observed that the obtained gaze data distribution was limited in x-direction since pilots' behavior inside the aircraft is unconstrained. We intend to collect data with large number of participants in non-flying outdoor conditions under high illumination and wider gaze directions. Such dataset collected in high illumination non-flying conditions can be used to augment the existing dataset. We shall also incorporate inter-person variations like IPD while designing our dataset and the gaze estimation model. We believe such dataset along with proposed machine learning model can help to obtain gaze estimates in actual flying conditions with higher accuracy and consistency.

6. SUMMARY

We presented the failure mode analysis of existing commercially available eye gaze trackers in actual flying environments. We demonstrated that the COTS eye tracker failed to work in higher illumination conditions using two flight recordings obtained in two fighter planes. We also demonstrated that the vertical tracking range of existing COTS eye tracker was not sufficient for aviation applications. We presented a machine-learning based approach for an end to end gaze estimation without relying on any handcrafted features. We performed within person and cross-person evaluations using the data obtained from two pilots. We demonstrated that our proposed architecture produced accurate gaze predictions with training data as minimum as 3 minutes recording for each pilot. Further, we demonstrated that the proposed architecture could produce gaze estimates reliably even when the COTS eye tracker failed to detect eye gaze.

ACKNOWLEDGEMENTS

We sincerely thank the pilots who took part in our experiment.

REFERENCES

[1] V. K. Sharma, L. R. D. Murthy, K. Singh Saluja, V. Mollyn, G. Sharma, and P. Biswas, "Webcam controlled robotic arm for persons with SSMI," *Technology and Disability*, no. Preprint, pp. 1–19, 2020.

[2] S. Eraslan, Y. Yesilada, and S. Harper, "Eye Tracking Scanpath Analysis Techniques on Web Pages: A Survey, Evaluation and Comparison," *Journal of Eye Movement Research*, vol. 9, no. 1, pp. 1–19, 2016, doi: 10.16910/jemr.9.1.2.

[3] V. Thoma and J. Dodd, "Web usability and eyetracking," in *Eye Movement Research*, Springer, 2019, pp. 883–927.

[4] G. Prabhakar, A. Mukhopadhyay, L. Murthy, M. Modiksha, D. Sachin, and P. Biswas, "Cognitive load estimation using ocular parameters in automotive," *Transportation Engineering*, vol. 2, p. 100008, 2020.

[5] G. Prabhakar *et al.*, "Interactive gaze and finger controlled HUD for cars," *Journal on Multimodal User Interfaces*, vol. 14, no. 1, pp. 101–121, 2020.

[6] M. D. Babu, D. v JeevithaShree, G. Prabhakar, K. P. S. Saluja, A. Pashilkar, and P. Biswas, "Estimating pilots' cognitive load from ocular parameters through simulation and in-flight studies," *Journal of Eye Movement Research*, vol. 12, no. 3, 2019.

[7] "TobiiProGlasses2 Accuracy Evaluation." 2021. Accessed: Oct. 14, 2021. [Online]. Available: https://www.tobiipro.com/siteassets/tobii-pro/accuracy-and-precision-tests/tobii-pro-glasses-2-accuracy-and-precision-test-report.pdf

[8] "TobiiDynavox Technical Specifications." http://tdvox.web-downloads.s3.amazonaws.com/PCEye5/Documentation/TobiiDynavox_PCEye5_UserManual_v1-0-2_en-US_WEB_12008591.pdf (accessed Oct. 14, 2021).

[9] M. Tonsen, C. K. Baumann, and K. Dierkes, "A High-Level Description and Performance Evaluation of Pupil Invisible," *arXiv preprint arXiv:2009.00508*, 2020.

[10] E. D. Guestrin and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Transactions on biomedical engineering*, vol. 53, no. 6, pp. 1124–1133, 2006.

[11] B. Brousseau, J. Rose, and M. Eizenman, "Accurate model-based point of gaze estimation on mobile devices," *Vision*, vol. 2, no. 3, p. 35, 2018.

[12] L. R. D. Murthy *et al.*, "Evaluating Accuracy of Eye Gaze Controlled Interface in Military Aviation Environment," 2020. doi: 10.1109/AERO47225.2020.9172480.

BIOGRAPHY

***L. R. D. Murthy**, received a B.Tech degree in electrical and electronics engineering from SASTRA University, India in 2015. He is currently pursuing a PhD at the Intelligent Inclusive Interaction Design (I3D) lab in the Centre for Product Design and Manufacturing, Indian Institute of Science. His research interests include Eye Gaze estimation, Computer Vision, Speech Processing and Machine Learning.*

***Pradipta Biswas** is an Assistant Professor at the Centre for Product Design and Manufacturing and affiliated faculty at the Robert Bosch Centre for Cyber Physical Systems of Indian Institute of Science. His research focuses on user modelling and multimodal human-machine interaction for aviation and automotive environments and for assistive technology. He set up and lead the Interaction Design (I3D) Lab at CPDM, IISc. He is a Co-Chair of the IRG AVA and Focus Group on Smart TV at International Telecommunication Union. He is a member of the UKRI International Development Peer Review College, Society of Flight Test Engineers and was a professional member of the British Computer Society, Associate Fellow at the UK Higher Education Academy and Royal Society of Medicine.*