U.S. Department of Transportation

Federal Railroad Administration

Monitoring Engineer Fatigue (MEFA)

Office of Research, Development
and Technology
Washington, DC 20590

DOT/FRA/ORD-19/48

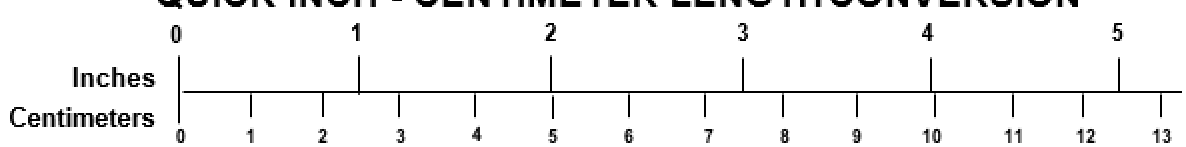
Final Report
December 2019

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.			
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 2019		3. REPORT TYPE AND DATES COVERED Technical Report, Sept. 2018 – March 2019
4. TITLE AND SUBTITLE Monitoring Engineer Fatigue (MEFA)			5. FUNDING NUMBERS 693JJ618C000015
6. AUTHOR(S) Corey Stein, Andrew Liu, Dan Brown, Amelia Porch			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Aurora Flight Sciences 90 Broadway, Suite 11 Cambridge, MA 02142			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Department of Transportation Federal Railroad Administration Office of Railroad Policy and Development Office of Research, Development and Technology Washington, DC 20590			10. SPONSORING/MONITORING AGENCY REPORT NUMBER DOT/FRA/ORD-19/48
11. SUPPLEMENTARY NOTES COR: Michael E. Jones			
12a. DISTRIBUTION/AVAILABILITY STATEMENT This document is available to the public through the [FRA eLibrary]().			12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) Aurora Flight Sciences, under contract to the Federal Railroad Administration, studied the feasibility of autonomously monitoring and combating train engineer fatigue. This program is called Monitoring Engineer Fatigue (MEFA). Research and testing at the FRA's Cab Technology Integration Laboratory confirmed that identifying fatigue through physiological cues through a variety of detection sensors is plausible. RGBD (color + depth) cameras were used to detect facial features for key points of interest – mainly eye and mouth movements. Researchers parsed an activity within the general code of operating rules (GCOR) into a logical step-by-step procedure, which could then be presented to Aurora's procedure execution module. The biggest area for improvement is the method by which fatigued engineers are analyzed. Researchers recommend future work include an analysis of engineers whom are actively fatigued. Relying instead on actors would increase the risk of developing a brittle algorithm.			
14. SUBJECT TERMS MEFA, engineer fatigue, monitoring, CTIL, Volpe, eye tracking			15. NUMBER OF PAGES 42 16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC			METRIC TO ENGLISH		
LENGTH (APPROXIMATE)			**LENGTH** (APPROXIMATE)		
1 inch (in)	=	2.5 centimeters (cm)	1 millimeter (mm)	=	0.04 inch (in)
1 foot (ft)	=	30 centimeters (cm)	1 centimeter (cm)	=	0.4 inch (in)
1 yard (yd)	=	0.9 meter (m)	1 meter (m)	=	3.3 feet (ft)
1 mile (mi)	=	1.6 kilometers (km)	1 meter (m)	=	1.1 yards (yd)
			1 kilometer (km)	=	0.6 mile (mi)
AREA (APPROXIMATE)			**AREA** (APPROXIMATE)		
1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)	1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)	1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)	1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)	10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)			
MASS - WEIGHT (APPROXIMATE)			**MASS - WEIGHT** (APPROXIMATE)		
1 ounce (oz)	=	28 grams (gm)	1 gram (gm)	=	0.036 ounce (oz)
1 pound (lb)	=	0.45 kilogram (kg)	1 kilogram (kg)	=	2.2 pounds (lb)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)	1 tonne (t)	=	1,000 kilograms (kg)
				=	1.1 short tons
VOLUME (APPROXIMATE)			**VOLUME** (APPROXIMATE)		
1 teaspoon (tsp)	=	5 milliliters (ml)	1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 tablespoon (tbsp)	=	15 milliliters (ml)	1 liter (l)	=	2.1 pints (pt)
1 fluid ounce (fl oz)	=	30 milliliters (ml)	1 liter (l)	=	1.06 quarts (qt)
1 cup (c)	=	0.24 liter (l)	1 liter (l)	=	0.26 gallon (gal)
1 pint (pt)	=	0.47 liter (l)			
1 quart (qt)	=	0.96 liter (l)			
1 gallon (gal)	=	3.8 liters (l)			
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)	1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)	1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)
TEMPERATURE (EXACT)			**TEMPERATURE** (EXACT)		
$[(x-32)(5/9)]$ °F	=	y °C	$[(9/5)\,y + 32]$ °C	=	x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO

°F	-40°	-22°	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 8/17/98

Acknowledgements

Aurora Flight Sciences would like to acknowledge Andy Liu from the Human Systems Laboratory at MIT. His research and report contributions allowed for a comprehensive understanding of the current state of human tracking algorithms. His efforts are much appreciated.

Aurora would also like to acknowledge the Federal Railroad Administration for allowing the research team to utilize its Cab Technology Integration Laboratory. With assistance from the U.S. Department of Transportation Volpe National Transportation Systems Center, the team obtained the necessary data to complete this program.

Contents

Executive Summary 1

1. Introduction 2
   1.1 Background and Motivation 3
   1.2 Objectives 4
   1.3 Overall Approach 4
   1.4 Scope 5
   1.5 Organization of the Report 5
2. Phase I 7
   2.1 Current State of Human Feature Tracking Algorithms 7
   2.2 Motion Sensor Down-select 14
   2.3 Ergonomics Analysis 15
   2.4 FRA CTIL Web Service 16
   2.5 Critical Data 17
   2.6 Hardware Systems and Software Diagram 17
   2.7 Representative GCOR Process Selection 20
3. Results 22
   3.1 Physical Testing Results 22
   3.2 Reason for Discontinuation 23
4. Conclusion 25
5. References 26

Appendix A. Pressure Sensor Images 29

Appendix B. List of CTIL Variables 33

Appendix C. Yellow Flag Procedure DTR File 34

Abbreviations and Acronyms 38

Illustrations

Figure 1: Aurora's ALIAS 2
Figure 2: Alerter System vs. MEFA 3
Figure 3: MEFA System 5
Figure 4: Intel RealSense D415 Camera 14
Figure 5: Intel RealSense D435 Camera 15
Figure 6: CAD Mock-up of the Train Cabin 15
Figure 7: FRA CTIL 16
Figure 8: MEFA Software Flow Diagram 18
Figure 9: Representation of Section 5.4.2 B in the GCOR Document, "Display of Yellow Flag, Restriction is Not in Effect" 20
Figure 10: Video Captured from Center Console – Standing Position 22
Figure 11: Video Captured from Center Console - Sitting Position 23
Figure 12: Video with Eye-tracking Software Implemented 23

Executive Summary

Aurora Flight Sciences, under contract with the Federal Railroad Administration, studied the feasibility of autonomously monitoring fatigue levels of train engineers. This program is called Monitoring Engineer Fatigue (MEFA). The intention was to utilize a vision sensor system to identify physiological cues of the engineer that could indicate fatigue. The system would then warn the engineer of potential danger based on the situation. The program period of performance was September 2018 to March 2019.

To monitor engineer fatigue, researchers took a multi-pronged approach. First, they used a visual sensor system to measure eye and mouth position, aspect ratio, and body position. Secondly, they designed a locomotive position/information state machine, using generic train operator processes. The combination of these two systems could then be used to observe the actions and fatigue level of the engineer and compare them with expected actions.

The research team confirmed that identifying fatigue through physiological cues through a variety of detection sensors is plausible. It determined that depth-sensing RGBD (color + depth) cameras would be used to detect facial features for key points of interest – mainly eye and mouth movements – as well as body and hand pose. Fatigue levels can be deduced based on eyelid blinking rates as well as an eye/mouth aspect ratio calculation. These cameras were installed in the Cab Technology Integration Laboratory at the Volpe Center for testing.

Cameras installed in a variety of locations captured video of a face and body from different angles. Facial feature tracking was completed using these videos with a 68-point facial feature tracker, which can identify the eyes and mouth on a face and calculate the relevant eye/mouth aspect ratios. Researchers also considered body and hand pose to determine possible signs of fatigue, such as slouching or non-movement for long periods of time. They also understood the importance of the operator's hands within the cab controls to judge possible actions being taken or not taken.

A functional locomotive position/information state machine was dependent upon the identification of an activity within the general code of operating rules (GCOR) that could be parsed into a logical step-by-step procedure. This is required to compare the engineer's expected actions with the measured actions. This system is a modified version of an existing Aurora design and has been tested vigorously in previous programs.

While this one procedure provides a great example of how this system would operate, in future work the locomotive position/information state machine will require many more activities from the GCOR and other sources to allow the module to gain an understanding of engineer activities. Additionally, this program had planned on utilizing test participants who would *act* fatigued, but the research team recommends the scope of future studies include test participants who are sufficiently fatigued such that they demonstrate ecologically valid physiological (e.g., eye closures, droopy head or body posture) and operationally erroneous behavioral cues (e.g., accelerating instead of decelerating). Based on the evidence in the literature, the team concluded that relying on the use of actors would increase the risk of developing a brittle algorithm that could not be reliably deployed in the field with confidence in its accuracy.

1. Introduction

This report details the research, development, and testing that Aurora Flight Sciences completed under contract to the Federal Railroad Administration (FRA) to determine the feasibility of combating fatigue in the railroad industry. This was accomplished by using a motion sensor system to identify physiological cues of the engineer that could be indications of fatigue (as measured directly, e.g., eye and mouth position, and indirectly, by comparing expected actions with actual engineer actions). The system would then warn the engineer of potential impending danger based on the situation.

This system would leverage modules of Aurora's Aircrew Labor In-cockpit Automation System (ALIAS), developed under DARPA. ALIAS (Figure 1) is an extensible, modular, and open toolkit designed to act as a digital "operator's assistant," providing another set of "eyes and ears" inside a complex vehicle environment. The Monitoring Engineer Fatigue (MEFA) system leverages two of ALIAS's five modules in conjunction with the commercial off-the-shelf (COTS) motion sensor to create a fatigue monitoring system: a) knowledge acquisition (KA), a means of digitalizing the general code of operating rules (GCOR) and capturing the task dependencies and parallels; and b) the perception system, which determines the state of the engineer using video sensing. Although originally designed for aircraft, ALIAS's emphasis on minimal invasiveness and rapid extensibility allows its core modules to be readily adapted for the rail industry, providing significant safety benefits. The intention behind MEFA is to provide a fail-safe technology that will reduce the number of accidents due to an intermediate fatigue state, where the engineer is neither fully awake nor fully asleep, and addresses an operational gap that current alerter systems cannot fill due to their design limitations.

Figure 1: Aurora's ALIAS

1.1 Background

Degraded performance due to fatigue is a significant contributor to major accidents in the rail industry, with engineers missing wayside signals, having reduced situation awareness, or falling asleep while operating the train. Every train engineer has experienced fatigue, routinely putting themselves, passengers, and areas neighboring the rail line at risk for accidents (Oman & Liu, 2006). The most common system used to keep locomotive crew engaged is the alerter system, which utilizes a reset timer to verify engineers are still awake and responsive. If an engineer does not respond to the alerter in time, the system will apply brakes to the locomotive. Although the current locomotive alerter system has contributed to keeping the engineer awake, nothing can determine whether the engineer is truly mentally engaged, as it is deactivated upon any interaction with the control system. Engineers may still be awake enough to press a button yet fatigued – lacking situation awareness of their surroundings or current operations. This gray area between fully asleep and fully awake is the need that MEFA aims to fill. A comparison between the alerter system and MEFA is shown in Figure 2.

		Technology	
		Alerter	*MEFA*
Mentally Engaged	*Physically Engaged* ***AWAKE***	× **Obtrusive**	✓ **No action**
	Physically Disengaged ***FATIGUED***	✓ **Detects**	✓ **Detects**
Mentally Disengaged	*Physically Engaged* ***FATIGUED WITH AUTOMATIC BEHAVIOR***	× **Cannot Detect**	✓ **Detects**
	Physically Disengaged ***FULLY ASLEEP***	✓ **Detects**	✓ **Detects**

Figure 2: Alerter System vs. MEFA

A deadly accident in Macdona, Texas in 2004 (NTSB/RAR-06/03) involved an engineer demonstrating automatic behavior, defined as a state where one is mentally fatigued but physically awake enough to continue providing input to the train control system. Because motor reflex responses typically require a lower level of cognitive effort (Abd-Elfattah et al., 2015), the operator could operate the train despite his impairment, which subsequently reset the alerter; it never triggered to rouse the engineer to a more alert state. However, the report also states that the engineer's control interactions were inappropriate given the task context, as the engineer increased the speed of the train when he should have been decreasing it. If MEFA had been in use before this accident, the engineer would have received continued alerts of inappropriate cab interaction activity upon engaging the throttle in the wrong direction. While the use of Positive Train Control (PTC) would have also prevented this accident, there are still multiple dark areas or rail line parts (e.g., grade crossings) that do not employ PTC, and nationwide implementation may be delayed by several years. Local obstructions or rail infrastructure failures, which cannot be captured by autonomous systems such as PTC, require immediate human intervention.

1.2 Objectives

The main objective of this program was to develop an in-cab passive monitoring system that detects and intervenes when a locomotive engineer is less vigilant due to fatigue. Researchers measured fatigue without relying on traditional, brittle eye-tracking systems. By the end of the program, the objective was to demonstrate feasibility and proof of concept in the FRA Cab Technology Integration Laboratory (CTIL) using realistic test cases with train engineers.

1.3 Overall Approach

The MEFA approach consisted of two major components: measurement and classification, each developed over three phases, with a total planned project time of 12 months. The final MEFA system is physically comprised of vision sensors, a laptop, and a tablet. The vision sensors measure the state of the cabin and the engineer. The laptop, used in the base MEFA program, synthesizes the various data streams and houses the fatigue classification algorithms. The not-safety-critical tablet, planned to be developed in an optional phase of this study, presents the required cab activity and simple feedback on engineer fatigue. Due to the flexible nature of this system, other fatigue intervention mechanisms could be used in conjunction with the classification component of MEFA. Later phases of the project would culminate in a proof-of-concept demonstration in the CTIL at the Volpe National Transportation Systems center and a final report on the research findings.

The MEFA system (Figure 3) is engaged at the start of the work shift. The vision sensor samples the engineer's head and body, and the data bus indicates the cab activity continuously through the entire shift. The different data streams are filtered and fused in real time. The classification algorithm takes these data and based on previous thresholds of universal behavior for fatigue, classifies the fatigue state. Concurrently, the system monitors and determines if the engineer is performing the expected procedures. This data is fused with the fatigue state to determine an overall analysis of the engineer's alertness. Finally, the fatigue state is forwarded to an intervention mechanism, such as a tablet, that activates if the engineer's state is inappropriate.

Figure 3: MEFA System

1.4 Scope

This system relies on multiple physiological and behavioral cues to infer the engineer's level of fatigue. These cues come from a few separate sources and are measured using COTS technology and an intelligent understanding of the engineer's activity in the cab.

MEFA synthesizes data from two main sources: the engineer's head and hands, and cab activity. Using these disparate cues provides redundancy and increases confidence in the accuracy of the fatigue classification. Furthermore, independent cue sources increase the robustness of the system to various working conditions that traditional fatigue-monitoring techniques cannot handle well: poor or overly illuminous lighting conditions, head apparel, eyeglasses, and excessive movement. This system is cue-dependent and sensor-independent, meaning that as sensing technology improves, the sensors themselves can be easily upgraded with minor modifications to the entire system.

1.5 Organization of the Report

Section 2 details the work completed during this program.

Section 3 includes the CTIL images taken by its sensors and discusses how this program will impact future work in this area.

Section 4 is a brief conclusion.

Appendix A shows pressure sensor images from this research on the current state of human-feature-tracking algorithms.

Appendix B lists many important variables accessible through the CTIL.

Appendix C is the full digital type-rating (DTR) file of the yellow flag procedure that researchers selected as the identifying procedure for fatigue detection.

2. Phase I

Phase I can be summarized by five sub-phases. Some of these happened concurrently, while others required results from previous sections to build upon. Section 2.1 details a literature review to understand the state-of-the-art in devices, physiological cues, and algorithmic approaches for detecting and tracking human fatigue. Section 2.2 briefly reviews the motion sensor selection process. Section 2.3 discusses the connection process with the CTIL. Section 2.4 describes the important data required for a successful system. Section 2.5 presents a hypothetical logical procedure that MEFA would be trained to examine.

2.1 Current State of Human Feature Tracking Algorithms

This report provides a brief list of physiological cues that have been used as input data for fatigue detection systems. The research and commercial literature were gathered from online search engines (e.g., Google Scholar), publisher websites (e.g., HFES), government databases (e.g., FRA), and company websites. First, three papers reviewing fatigue detection technology (one for rail, two for automotive) are briefly summarized, but the individual systems that were reviewed are discussed under the discussion of individual cues.

2.1.1 Reviews of Fatigue Detection Devices

The most recent review of fatigue detection devices specifically for the rail industry was a report to the UK Railway Safety and Standards Board in 2002 [1]. It was a comprehensive review of then-current devices measuring alertness and readiness and evaluated their feasibility for use on the UK Rail Network. Eight systems were based primarily on ocular measures (e.g., eye closure, blink rate), and while four other devices utilized skin conductance, head movements, EEG, or secondary task response time. Only two of the reviewed devices were non-eye movement based system deemed as suitable for rail use: (1) The Engine Driver Vigilance Telemetric Control System (EDVTCS),[1] which is an electrodermal-based system, and (2) the Micro-Nod Detection System, [2] which detects patterns of head nodding.

No other reviews aimed at rail operations have been found, but several reviews of fatigue detection systems for automobile and truck applications have been performed [3–6] since the RSSB report. The two most recent are discussed here. In 2014, Dawson et al. [3] examined 16 different systems that performed continuous operator monitoring grouped by the primary method of measurement. Five systems used eye-based measures, two systems were based on EEG, seven were based on posture or head nodding, and one system was based on galvanic skin resistance. At the time, the authors concluded that "… none were sufficiently well validated to provide a comprehensive solution to managing fatigue-related risk at the individual level in real time." Since then, several eye-based systems have now become commercially available. The Seeing Machines Driver Safety System is presently commercially available for both rail (with

[1] http://www.neurocom.ru/en2/product/edvtcs.html (last accessed 18 January 2019)

Progress Rail[2]) and automotive (Guardian System[3]) applications. SmartEye has claimed[4] that "German premium car manufacturers" and Geely (Chinese OEM) have installed their technology in current vehicles. The Optalert system has been tested in aviation, mining and road transport, but there is no mention of current commercial use on the company website.[5] The Smart Cap EEG-based system[6] is also now commercially available but requires a sensor band worn on the operator's head, usually beneath a cap or helmet. Two case studies from the mining industry (truck drivers) claim a reduction in fatigue-related incidents. Head nodding detectors are widely available on Amazon.com for as low as $5 but the authors point out that these systems only detect sleep episodes after the fact and with low specificity and require a device to be worn on the head. The Micro-Nod system, also cited in the RSSB report, has not been developed into a commercial product. The EDVTCS, also mentioned in the RSSB report, appears to still be commercially available although no results from field trials has been found.

Sikander and Anwar [6] provide the most recent technology review (2018) and briefly describe additional existing systems. Automobile systems from Toyota, Nissan, and Volkswagen assess driver behavior (e.g., steering inputs) as the observable correlate of sleepiness, although Toyota and Nissan originally employed camera-based eye detection systems [7]. Additional commercial products found by Sikander and Anwar include the Care Drive's MR688[7] system and the GuardVant OpGuard[8] system, both which use single IR-camera systems to detect eyelid closure, pupil changes, and head movements. The authors also provide a wide survey of projects using multiple cue combinations for fatigue detection. Other commercial systems for the automotive industry can be found in a New York Times article "Sleepy Behind the Wheel? Some Cars Can Tell."[9]

2.1.2 Physiologic Cues Measured by Fatigue Detection Systems

Eye-based measures

The most commonly employed oculomotor measure for fatigue detection systems is eyelid closure which was found to be a reliable predictor of the onset of sleep [8, 9]. PERCLOS – the percentage of time of eye closure – is commonly defined as the percentage of a time interval (typically 3 - 6 mins) that the eyes are significantly closed (e.g., 80 – 100 percent). Early IR-camera based fatigue detection system using PERCLOS measurement (e.g., Attention Technology DD- 850 system ("Co-Pilot"), were not commercially successful due to problematic data loss stemming from failures to track head motion [10], misinterpreted glances

[2] https://www.progressrail.com/en/innovation/fatigue.html (last accessed 5 February 2019)

[3] https://www.seeingmachines.com/guardian/ (last accessed 5 February 2019)

[4] https://smarteye.se/automotive-solutions/(last accessed 5 February 2019)

[5] http://www.optalert.com/ (last accessed 4 February 2019)

[6] http://www.smartcaptech.com/life-smart-cap/ (last accessed 5 February 2019)

[7] http://www.care-drive.com/product/driver-fatigue-monitor-mr688/ (last accessed 5 February 2019)

[8] https://www.guardvant.com/how-opguard-helps-drivers-combat-fatigue/ (last accessed 5 February 2019)

[9] https://www.nytimes.com/2017/03/16/automobiles/wheels/drowsy-driving-technology.html (last accessed 17 February 2019)

(e.g., toward the dashboard) or occlusion from eyewear [11]. Eyelid closure is also not consistently observed in subjects who exhibit sleep patterns in EEG (described in [12]). Indeed, subsequent work by Wierwille and colleagues [11, 13] focused on a combination of PERCLOS with driving performance measures such as lane-keeping. Commercial aftermarket fatigue detection systems, which do not have access to driving performance parameters, also use multiple oculomotor measures such as PERCLOS, pupil measures, and head movements.

Johns et al. developed a commercial IR reflectance oculography system (Optalert) that measures several non-PERCLOS oculomotor variables, such as slow-rolling eye movements or changes in maximal saccadic velocity of eyelid re-opening, and combines them into a single value on the proprietary Johns Drowsiness Scale (JDS) [14]. The hardware system is built into the frame of a pair of glasses which must be worn by the user, and it is claimed that the JDS does not require any recalibration for individuals.

Schleicher et al. [12] examined the correlation of several oculomotor parameters such as blink frequency and duration, saccadic velocity, fixation duration to subjective ratings of sleepiness. From their data collected in a driving simulator, the three oculomotor parameters with the highest correlations were blink duration, delay in lid reopening, and lid closure speed. The authors note two important points in their discussion: (1) Obtaining continuous high quality measurements in an operational system may be problematic and (2) Individual differences in the distributions of oculomotor parameters will make selection of detection threshold more difficult. Boverie and Girard [15] developed a fuzzy classifier based on these parameters to classify driver state into one of four classes ("sleepy," "drowsy," "slightly drowsy," "alert"). In testing on real-world driving data, the average sensitivity was 86 percent, while specificity was 89 percent.

Several other eye parameters have been explored as indicators of drowsiness. Catalbas et al. [16] collected data on the frequency of peak saccadic accelerations above a threshold for "vigorous" and "fatigued" drivers (neither term was defined) and, through an offline analysis, determined a difference between the mean and standard deviations of the peak acceleration distributions. However, they did not use any statistical testing to determine if the differences were statistically significant. Although saccadic eye movement parameters had a relatively low correlation to subjective sleepiness in [12], peak velocity of saccades has been shown to be sensitive to sleep deprivation [17] and DiStasi et al. [18] found that similar changes could be detected after 2 hours of driving in a simulation, suggesting that peak velocity could be used to detect the onset of fatigue.

Xu et al. [19] used average fixation time and pupil area from 1- to 2-hour driving simulator sessions to train a fuzzy K-nearest neighbor classifier. The fatigue state was assessed on a subjective scale and overall classification sensitivity was almost 89 percent. A classifier trained with both cues had better performance for most subjects while the single-cue trained classifiers were slightly less accurate but similar to each other in performance. Deng et al. [20] noted that the frequency of pupil diameter changes were greater in subjects after sleep deprivation (<4 hours sleep). Schumann et al. [21] similarly found that pupil size variability, measured by standard deviation, Shannon entropy, and wavelength entropy, was higher just before sleep onset. They were able to correctly identify 83 percent of the pre-sleep data segments (30 sec. duration) using linear discriminant analysis of all three variables.

Head pose/movement

As noted above, several inexpensive commercial fatigue detectors detect high-acceleration head movements forward or head angle as an indicator of microsleep episodes or drowsiness, but these provide alerting only after an event has occurred. The Micro-Nod Detection System [2] attempted to classify patterns of head movements through three non-contact capacitive sensors mounted above the head in the car ceiling headliner with PERCLOS used as ground truth for fatigue state. A 15- to 20-minute training period was needed to establish "normative" head behavior for an individual driver. If the observed behavior exceeded normal (threshold not defined in report), the system would trigger a visual alert. No details about how the measured capacitance was processed to label processed head movements. The author claimed that drowsiness detection in advance of a driver error (e.g., road departure) using this system was higher than the performance of a generic PERCLOS system. There is no evidence that Micro-Nod system was ever developed into a commercial product or tested in a rail or automobile system.

Body pose/movement

The EU project known as SENSATION[10] aimed to develop new sensors to detect operator state in real time. One measurement concept was a pressure-sensitive sensor foil (SEFO) to assess body posture and movement. This approach was described in Mora et al., [22] where the sensor mat was embedded in a hospital bed to detect respiratory events indicating sleep apnea. Although recognition performance was high in this context, the authors noted that performance would be difficult to reproduce in the dynamic environment of a moving vehicle. Nakane et al. [23] investigated whether **postural sway** in a seated position would converge more quickly to a baseline location when subjects were more likely to become fatigued (evening test time) than not (morning test time). Their experiment results showed that the center of pressure for the left-right axis for subjects experiencing fatigue converged earlier than for those experiencing non-fatigue. Subjects in the study sat quietly for 2 minutes. Chen [24] proposed a machine-learning method to classify automobile driver fatigue based on **21 features** (see Appendix A) from the seat and backrest pressure distribution and acceleration data from the shoulder belt. Using data collected from 2-hour driving simulator sessions performed in the morning (alert state) and mid-afternoon (fatigued state), while fatigue was subjectively rated in 5-minute sampling periods both by the subject (1–100 scale) and by a rater using video of the driving session. Support vector machine (SVM) and hidden Markov model (HMM)-based fatigue classifiers performed with 88 percent and 90 percent accuracy detecting five levels of fatigue, although performance was relatively better for the three lowest levels of fatigue compared to the two highest. Further research is needed to determine whether these methods would be effective in the dynamic environment of actual rail operations, where extraneous vibrations could add considerable noise to the pressure measurements and affect recognition accuracy.

Facial recognition

Machine-vision techniques are now being applied to fatigue and drowsiness detection by detecting combinations of facial features such as **yawning** and **eye closures**. As one example, Zhang and Hua [25] have taken one approach of using a SVM classifier to learn facial features of

[10] SENSATION http://www.sensation-eu.org/

the drowsy and non-drowsy state and distinguish between them. Using a single camera set-up, their system achieved about 85 percent recognition accuracy with the driver simulating the effects of fatigue on facial and eye behavior during an on-road test. Other efforts have looked at automatic detection of specific facial expressions such a yawning, using techniques such as convolutional neural networks [26] or linear discriminant analysis [27, 28]. Using previously collected on-road video data, Zhang et al. [26] achieved a 92 percent detection rate (13 percent false positives) with yawning events present.

ECG, EMG, skin conductance

Non-contact measurement of these physiological signals would be the most desirable, but bodily movements present a problem ensuring reliable data collection. Wearable devices, such as a wrist-worn fitness device or smart watch may provide more consistent data collection, but require the operator to remember to don the device. Embedding measurement devices into the physical controls is a third option, but train control requires only episodic interaction, unlike continuous steering needed to control a vehicle, the data from these sensors would be highly fragmented.

Solaz et al. [29] tested the efficacy of detecting respiration rate from chest movements using a video camera system. They demonstrated that chest movements detected with the camera system was highly correlated to motions recorded with a plethysmography band. However, they had not progressed to developing a system to automatically detect respiration rate or fatigue that could be tested in a driving simulation. Wusk and Gabler [30] demonstrated the feasibility of using a car seat sensor designed for occupant classification of the occupant's respiration rate (RR) and heart rate (HR) in a laboratory setting. The system was within 10 percent of true RR and HR values with their data processing techniques, but further testing under real-world conditions is needed to determine whether the sensor will have the necessary sensitivity to operate in real-world situations. Yu [31] developed a non-contact electrocardiogram (ECG) detector to obtain respiratory and eye blink data, which was evaluated in a high-fidelity car simulator. Subjects who had been sleep-deprived were then evaluated before and after a 15-minute driving session. They found small differences in average values or HR, HR variability, and blink rate, but a fatigue assessment metric was not developed based on these metrics. The authors also reported performing a field test but did not provide any quantitative information about performance. Plessey Electronics has developed a system called WARDEN,[11] which uses seat-based sensors to detect ECG signals which are used to calculate heart rate variability. Details were not included in the product literature but the system appears to compute a normative HRV range and detects any exceedances of the range which then require the driver to acknowledge a visual alert on a separate display.

Rogado et al. [32] used a steering wheel-based sensor to detect EKG and grip pressure during a non-driving, steering-like task. The heart rate variability (HRV) signal was the measured interval between heart beats, but the authors also computed the fast Fourier transform (FFT) of this signal and compared the power in low- and high-frequency bands for data collected when subjects were alert or sleep deprived (not defined in paper). The authors noted a transient change in HRV when subjects fell asleep during the task, accompanied by spikes in the power spectrum.

[11] http://www.plesseysemiconductors.com/products/warden/ (last accessed 18 Jan 2019)

Fu and Wang [33] proposed a fatigue detection system using a "non-contact" sensor to measure electromyograph (EMG) and ECG but subjects wore an arm sleeve with detectors (as opposed to using skin surface electrodes) and had a reference electrode attached to the ankle. Periods of subjective driver sleepiness during their driving session were indicated by the subject raising their hand. Offline analysis of the data indicated that the best features for recognition were peak EMG amplitude and the maximum of the cross-correlation between EMG and ECG signals collected over 30-second epochs.

The Engine Driver Vigilance Telemetric Control System (EDVTCS) is a wrist-worn electrodermal sensor which had been in service on Russian railroads. A "relaxed" state is calibrated at the start of the trip and an unspecified algorithm purports to detect when the driver loses alertness and concentration and triggers an alarm which must be reset before the brakes are automatically applied. Dorrian et al. [34] tested the device under laboratory conditions and concluded that the system was not sensitive to sleepiness and fatigue induced by sustained wakefulness (28 hours) in their study. Lee et al. [35] proposed a similar wrist-worn system using both skin conductance (GSR) and hand and forearm EMG to detect fatigue in automobile drivers. For actual testing, the authors used surface electrodes to measure EMG and fingertip sensors to measure skin conductance rather than the proposed wrist-worn detector. They calculated both time-domain features (i.e., mean, variance kurtosis, and skewness of the signal amplitudes) and frequency-domain features calculated from a wavelet packet transform (WPT) (i.e., signal power at peak, mean, median, low, and high frequencies, low/high ratio[12]). These features were used in a support vector machine (SVM) classifier trained on 70 percent of the collected data.

Using the remaining 30 percent of the data for testing, the system showed 88 to 94 percent accuracy classifying the level of fatigue specified by subjects using the Karolinska Sleepiness Scale.

Khushaba et al. [36] developed a set of fuzzy-classifier of wavelet-based features using data from EEG, EOG, and ECG, and compared classification performance with subjective raters evaluating video recordings of the subjects' faces using criteria based on Wierwille and Ellsworth [37]. EEG, EOG, and ECG signals were collected using skin electrodes during two 25-minute driving sessions performed between 9 a.m. and 1 p.m. The first session was in higher-density traffic environment, the second in a low-density monotonous environment. Instead of using group statistics of frequency or power derived from FFTs of the signals, they instead used the coefficients from the WPT of the signals to construct feature sets for the classifier. They achieved recognition accuracy of 95 to 97 percent with a combination of all three data types. Accuracy fell to as low as 80 percent with fewer input signals.

Engineer control movements

Engineer control behavior or train state could be used as additional evidence to identify or predict fatigue. For example, Dorrian et al. [38] examined the effects of fatigue on the ability to negotiate speed restrictions. They found that drivers in a high-fatigue group used the brake less and traveled at higher speeds on certain terrain (e.g., descending grade) but not across all terrain conditions. In two of the four speed restrictions, average train speed was in excess of the posted speed. Behavioral measurements such as this could be integrated into a computational

[12] The terms "low" and "high" were not defined in the paper. Presumably they may refer to the average frequency within a low and high band.

approach to infer engineer state, similar to efforts to incorporate driver behavior into fatigue detection [39]. In the automobile industry, this approach has been incorporated with in-vehicle control behaviors such as steering wheel movements [40]. The behavior provides some context for what actions should occur and when they should occur. This type of context information has been shown to reduce false alarm rates compared to a PERCLOS-based fatigue detection system [41]. Mercedes Attention Assist[13] primarily measures small changes in driver steering behavior to gauge fatigue. It calibrates a baseline pattern from the initial driving behavior on a trip and compares current behavior in real time. Driving conditions, other activity such as interacting with the controls, and car state are also considered before an alert is issued.

2.1.3 Recommendations

The automotive market is currently the strongest driver for the development of fatigue detection technologies. As noted above, two general approaches have been taken for commercial products. Automotive manufacturers are developing systems that primarily use car-based information, including driver control behavior, to infer the driver's state. Third-party developers primarily rely on eye-based measures of driver fatigue or drowsiness. If these systems are integrated into a car, they may also access car-based information to improve recognition performance. There is little evidence that other physiologic sensors such as HRV, EMG, or seat pressure distribution – which could potentially be measured non-invasively – have the required accuracy and stability to be used in an operational environment.

The rail operating environment shares some similarities with the automotive environment in that a single, seated operator interacts with various controls to operate the vehicle. But there are notable differences that potentially make either automotive eye-based or behavior-based detection systems more difficult to adapt to the locomotive cab environment. The larger physical space of the locomotive cab, as Aboukhalil et al. [10] noted in their evaluation of the DD-850 infrared camera-based detection systems, means that any eye-based system would have to account for range of engineer's head and body motion in order to track their facial features and would as well as for periods when the operator is entirely out of view. Vehicle control behavior is also quite different. Continuous directional control of the train is not required, so there is no analog engineer action to steering movements. The dynamics of train handling occur on a longer time scale, so control actions such as changing the throttle settings tend to occur less frequently, although may occur rapidly around particular events, such as approaching a speed restriction. The effective "sample rate" of behavior may be too slow to catch changes in engineer state (e.g., fatigue) in a timely fashion. Finally, the temperature and vibration environment in the locomotive cab is much less controllable and hospitable, which may reduce the effectiveness of physiological measures such as GSR, ECG, or seat pressure distribution.

Nevertheless, a multiple camera-based approach to monitor engineer posture and activity combined with knowledge of the current operating scenario (e.g. speed limits or track properties which are available through the PTC system) is the recommended approach for developing a fatigue and drowsiness attention for locomotive engineers. Commercial off-the-shelf RGB-depth camera systems such as the Microsoft Kinect are widely available, are relatively inexpensive,

---

[13] https://www.mbusa.com/mercedes/technology/videos/detail/title-safety/videoId-710835ab8d127410VgnVCM100000ccec1e35RCRD (last accessed 17 February 2019)

can track human posture relatively accurately, and could potentially be modified to obtain other physiological parameters (e.g., Solaz et al. [29]). This approach provides simultaneous measurement of multiple cues from which to infer fatigue: overall activity level, head direction, head nodding, and body posture. Furthermore, the same system can be used to reconstruct the 3D environment of the cab, so the engineer's physical actions could also be inferred from the video stream (e.g., is the engineer moving the throttle or responding to the alerter?). This approach would enable the detection system to operate separately from the locomotive infrastructure but also supplement the functionality of integral systems such as the alerter, trip optimizer (TO), or PTC. If the system could access information from TO or PTC, such as the ideal notch or brake settings or upcoming speed or signal restrictions, then the timing of particular actions could be used to further characterize the state of the engineer (e.g., is the braking action occurring unusually later and with higher force compared to the engineer's typical response to an upcoming speed restriction?)

2.2 Motion Sensor Down-select

In parallel to the fatigue tracking literature review, researchers conducted a trade study of six potential monitoring sensor systems, exploring their suitability for use in the train cabin. These are the sensors, in no specific order:

- Intel RealSense
- Microsoft Kinect
- RGB Camera
- Leap Motion
- Optalert Glasses
- LUI (SixSafety Systems)

Considering the characteristics of each sensor and the recommendations from the fatigue tracking research, the team determined that the Intel RealSense was best-suited for this project, the primary drivers of which being that the IP ownership was open source, the camera was able to detect and classify facial emotions, and there was in-depth resolution and frame rate control of the camera. For purposes of this report, any discussions of video data recording should be assumed to be done by the Intel RealSense sensors acquired specifically for this program. The RealSense D435 (Figure 4) and D415 (Figure 5) were the models purchased.

Figure 4: Intel RealSense D415 Camera

Figure 5: Intel RealSense D435 Camera

2.3 Ergonomics Analysis

A number of camera mounting locations were possible within the CTIL train cabin. Aurora visited the CTIL in December 2018 to take measurements of the cabin and identify where the best mounting locations existed.

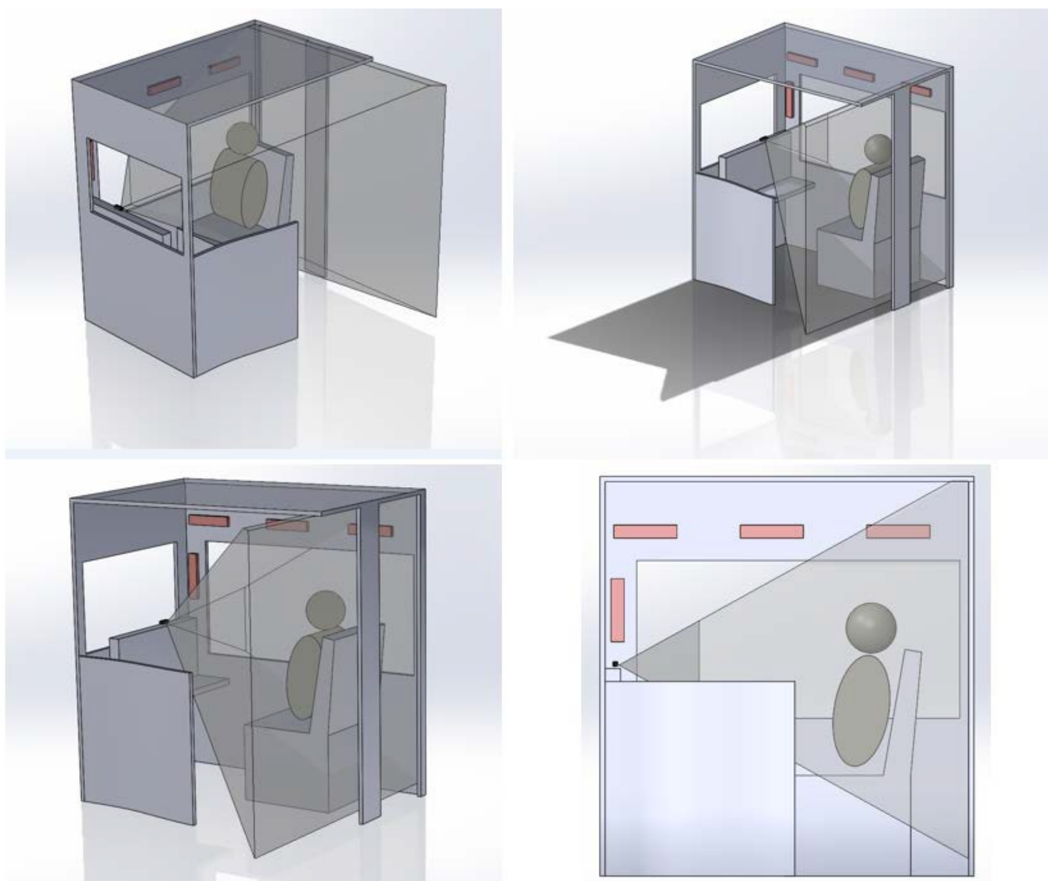


Figure 6: CAD Mock-up of the Train Cabin

The CAD mock-up (Figure 6) allowed the team at Aurora to test a variety of camera locations given the field of view (FOV) angle of the cameras acquired. Once these locations were identified, the cameras were taken to the CTIL to obtain recordings while the cameras were mounted in each position. The goal was to identify the best camera and location that would allow complete view of the engineer while he or she was sitting or standing. Results of this study are included at the end of this report.

2.4 FRA CTIL Web Service

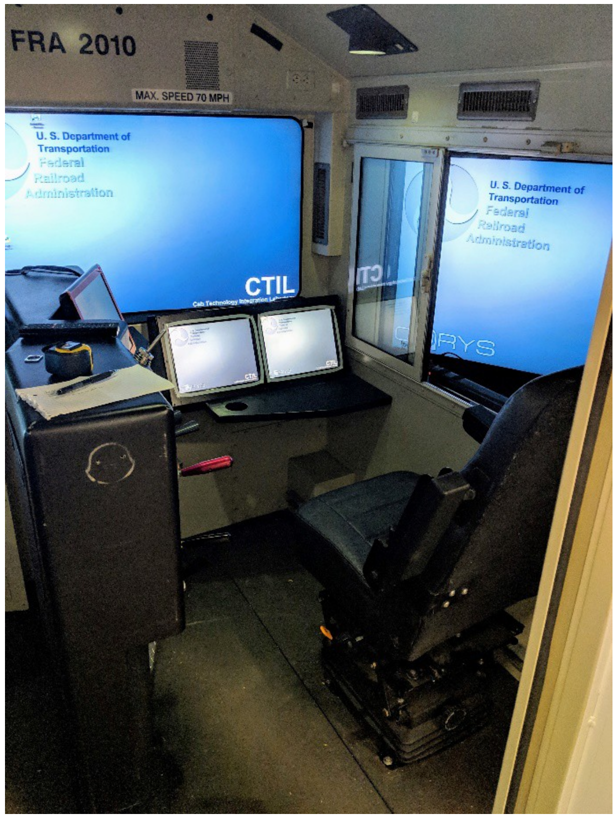


Figure 7: FRA CTIL

The FRA CTIL (Figure 7) can transmit and receive data over its proprietary web service data stream. The web service allows data to be sent around the CTIL such that any computer connected to the CTIL network has access to locomotive-related information. A few examples of such information available are position, speed, acceleration, grade, positions of all cab controls, brake pipe pressures, cylinder pressures, ER pressures, ammeter, fuel use variables. A list of these variables is included in Appendix B. While there are dozens of separate, obtainable values from the CTIL, the most important values for MEFA are related to understanding the state of the locomotive and how that relates to the operating engineer.

2.4.1 Connecting to the CTIL Web Service

Access to the CTIL data stream is established by a web service running on HTTP 1.1, connecting via port 80. A data retrieval request is formatted by a specific web URL which returns the requested information. Any computer connected to the CTILs network, either over wifi or Ethernet, has access to this data stream. The core of the URL request is:

http://192.168.0.2/data.csv

This core request is augmented with the specific parameters being asked of the web service. The request is formatted using the standard URL parameter syntax, "?" to start the parameter list and "&" to separate parameters. Below is a list of the parameters:

"v" – This required parameter allows one to specify simulator variables you'd like to receive. You can list multiple variables separated with commas. Any specified parameters that does not exist will simply be ignored.

"r" – This optional parameter specifies the data rate in hertz with witch the web service will respond. Its range is 1 to 20 Hz. If left unspecified the rate default rate will be 20 Hz.

"n" – This optional parameter specifies the number of responses the data stream will respond with. I If left unspecified the web service will default to n = 1, returning only receive one result.

"format=-h" – Used to specify a data return without a header row.

":|" – Optionally specify delimiter type. In this example the "|" delimiter is used. If left unspecified the default value is a "," (comma). An "&" is not required to separate this variable but does need to be the final parameter specified.

For example, to collect information on throttle and dynamic brake handle at 5 Hz for 20 seconds without a header row, with a "," as the delimiter. The web URL request will look like:

http://192.168.0.2/data.csv?v=thtld,dbhld&n=100&r=5&format=-h:,

To programmatically retrieve data via this web service, one can use the cURL library tool, which allows the transferring data with URL syntax. This data can be frequently requested over the web service to obtain the current state information about the CTIL.

2.5 Critical Data

Basic information for determining the current situational awareness of the driver would relate values such as current location, speed, acceleration, or braking information based on a set of predetermined limits for a particular segment of track. If the operator exceeds any of these limits, an alert would be sent to the operator via the monitoring system's user interface. More detailed information such as the position of the cab controls can also be specified limits based on a given track location. The CTIL web service would allow one to monitor the state of the alerter, alarm bells, bail status, percentage of the independent applied, direction of the reverser, and the speed notch position. This information can also be paired with computer vision acquired location of the operator's hands – indicating an intended control change or preparation for a control change.

2.6 Hardware Systems and Software Diagram

There are three major sections of the software diagram (Figure 8) for MEFA. Each provides a significant understanding to the state of the locomotive, the operator, and the two combined. These systems will be run off a single computer with one or two Intel Realsense depth cameras to provide video of the operator to the system. The computer will also be connected over Ethernet to the CTIL's network to retrieve data from the web service. Each section will be discussed in detail in the following sections: the three sections are locomotive position/information state machine (with information provided from the GCOR), the head/body/hand pose estimation, and the locomotive and operator fused situational awareness module.

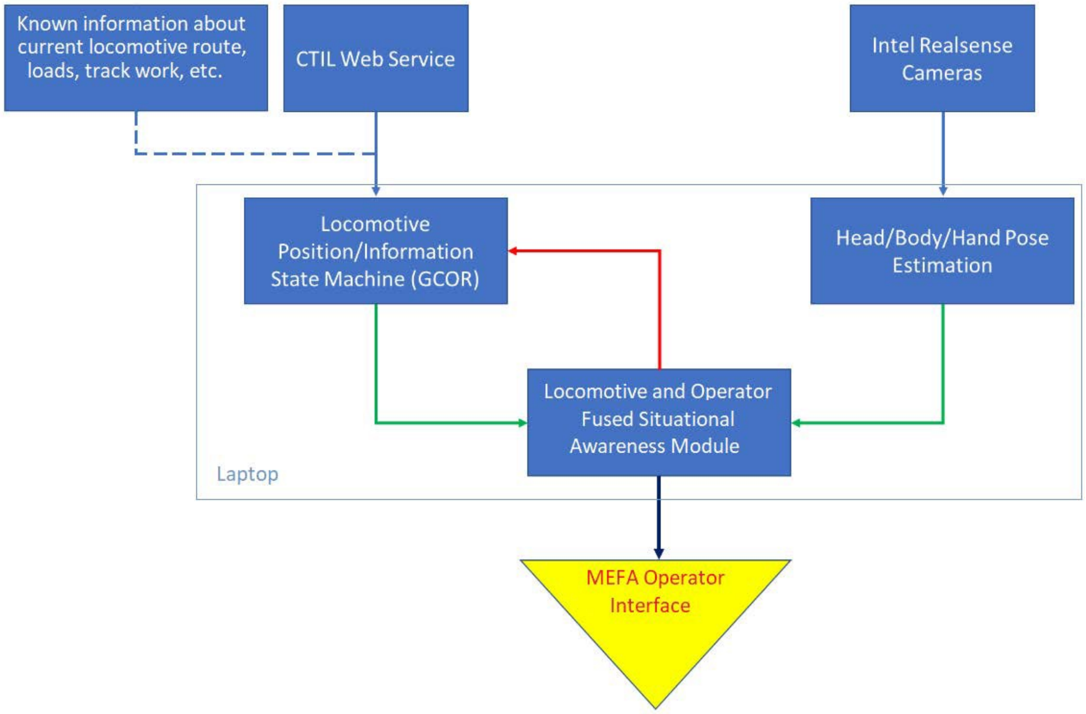


Figure 8: MEFA Software Flow Diagram

2.6.1 *Locomotive Position/Information State Machine (GCOR)*

Data obtained via the CTIL web service will inform this module of the current state of the locomotive. To gain a more comprehensive understanding of the current state of the locomotive (and therefore create a more robust system), this information on internal characteristics such as position, speed, acceleration, and braking pressures will be combined with known information of external characteristics from TO or PTC about the current locomotive route, loads, and any track work or bulletins for the entirety of the route. This information will be "truth-checked" against the locomotive position/information state machine based on the GCOR. With prior knowledge of how the train should operate based on the given location and load information, a "operation out-of-GCOR bounds" signal is passed onto the fused locomotive and situational awareness module. Note this fused locomotive and situational awareness module can also send a signal to the locomotive position/information module to preclude any warnings based on in-progress action taken to realign the locomotive's operation within the GCOR guideless.

2.6.2 *Head/Body/Hand Pose Estimation*

Depth-sensing RGBD (color + depth) cameras will be used to detect facial features for key points of interest, mainly eye and mouth movements. Findings from the fatigue tracking literature review (shown in Section 2.1), confirmed that relevant drowsiness factors can be deduced based on eyelid blinking rates as well as an eye/mouth aspect ratio calculation. Facial feature tracking will be done with the Dlib[14] implantation of a 68-point facial feature tracker.

This will allow researchers to find the eyes and mouth (along with many other features) and calculate the relevant eye/mouth aspect ratios.

Body and hand pose will also be used to determine any possible signs of fatigue, such as slouching or non-movement for long periods of time. More pertinent will be the hand detector, which will locate the operator's hands within the cab controls. It is important to localize the hands to judge possible actions to be taken or actions which might not be taken. This information along with any fatigue information from the facial feature tracking is passed onto the fused locomotive and situational awareness module.

2.6.3 Locomotive and Operator Fused Situational Awareness State

This module receives the individual pieces of current-state information from both the locomotive and the human operator. Based on thresholds for locomotive operation and any detected fatigue, a warning will tell to the operator to regain control. Along with this warning, this module will present recommended recourse to correct either the train operator to lie within the GCOR or inform the operator of his/her fatigue.

[14] [1] Davis E. King. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research* 10, pp. 1755-1758, 2009

2.7 Representative GCOR Process Selection

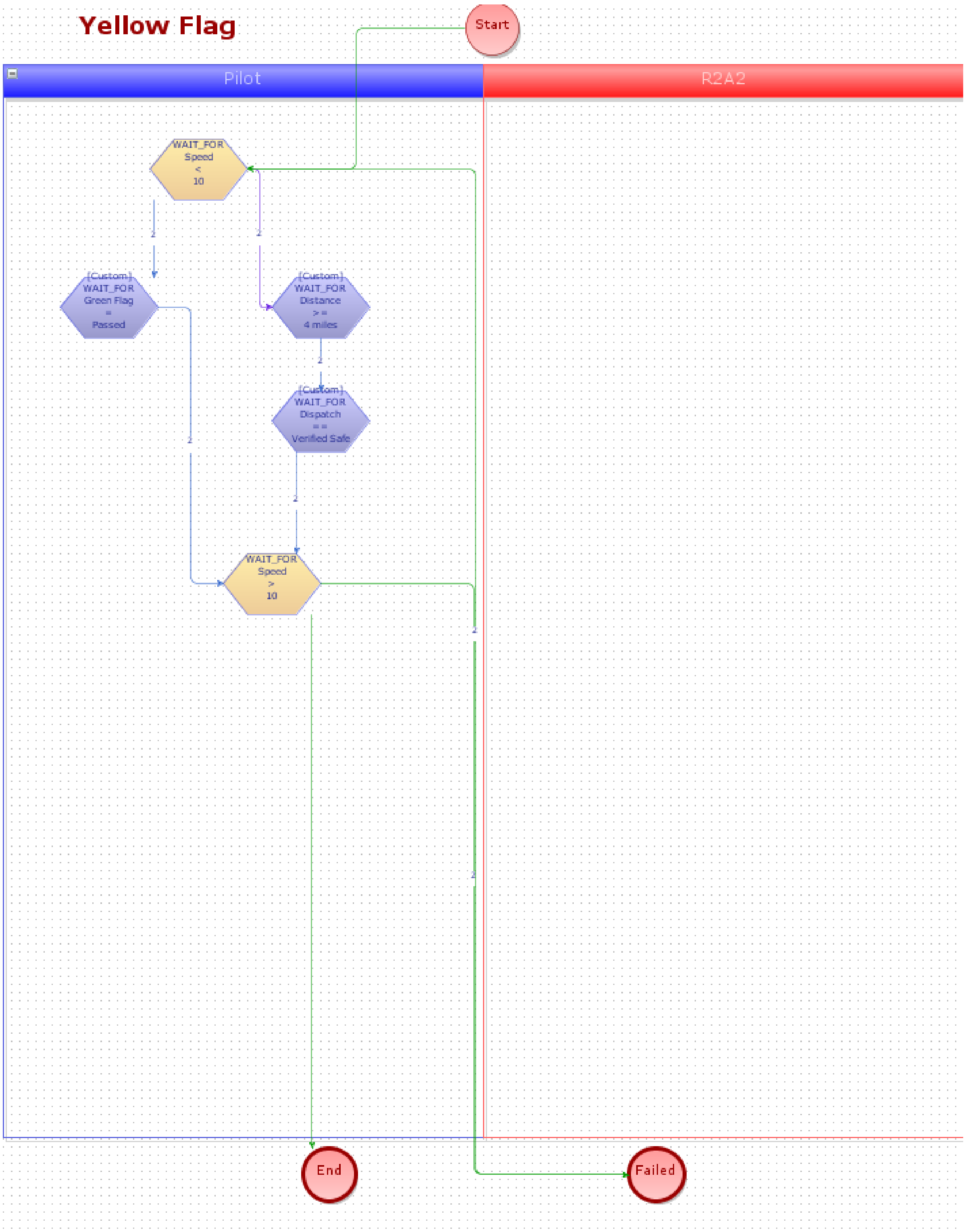


Figure 9: Representation of Section 5.4.2 B in the GCOR Document, "Display of Yellow Flag, Restriction is Not in Effect"

A process from the GCOR was identified as a representative procedure for the locomotive position/information state machine, and a logical step-by-step procedure was created. The GCOR process identified was the “Display of Yellow Flag” when a Restriction is Not in Effect. A yellow flag is displayed 2 miles before the restricted area due to track conditions or structures. However, if there is no restriction defined, the engineer must move the train at no faster than 10 mph. This can end once the engineer sees a green flag or has traveled over 4 miles beyond the yellow flag and a dispatcher has confirmed that there is no speed restriction in effect at that location.

A graphical representation of this process is shown in Figure 9. In this figure, each logical element of the process is represented by a physical block, and is represented in the locomotive position/information state machine as a “task.” If the block is yellow, it can be observed through existing CTIL data. If it is blue, the train operator will have to verify that the task has been completed. This can occur actively, by the system requiring the operator to press a button (or perhaps perform a gesture that is observable to the MEFA motion detection system in the cab). It could also occur passively, with the system assuming the task has completed and moving on to observe the next available block. Once a task is considered to be completed, the procedure will either move on to the next action or wait for one of several paths to become true if there are multiple options. Each block also has a timeout value. This means that if the task’s action has not occurred within the time allowed and the task has an output failure state, the task will time out and the system will considerer the procedure failed.

This procedure is expected to activate (either by operator activation or some external signal) once a yellow flag is observed with no restriction in effect. The system will use CTIL data to observe that the speed has dropped below 10 mph. Then the operator has two choices. If a green flag is observed, he or she can mark that block as complete. Otherwise, if the train has traveled 4 miles and the dispatcher has verified the route to be safe, the operator can mark the “Dispatch Verified Safe” block to be true. Finally, the system waits for the train to speed up, and ends the procedure. If either of the speed blocks take more than 60 seconds to activate, the procedure is considered to be failed.

This procedure will be presented to the procedure execution module using the “DTR” file (Appendix C), which represents the data types and procedure blocks present in the system.

3. Results

Phase I results are presented below, including CTIL testing and explanation of program discontinuation.

3.1 Physical Testing Results

Video was captured using the Intel RealSense D435 and D415 inside the CTIL cabin. There was confirmation that the center console position allowed for the best viewing angle of the face, whether the engineer was standing (Figure 10) or sitting (Figure 11), although all camera locations were evaluated. Other angles show the body of the engineer. The cameras can be mounted vertically or horizontally; however, due to the frame size of 1920 x 1080, it is advantageous to mount them vertically to get the most range within the image. Although both cameras provided adequate video quality, researchers identified the D435 as superior due to its wider FOV. Additionally, the D435 was able to mount in a smaller space due to its form factor, which could prove useful if the mounting location were to move in a future program.

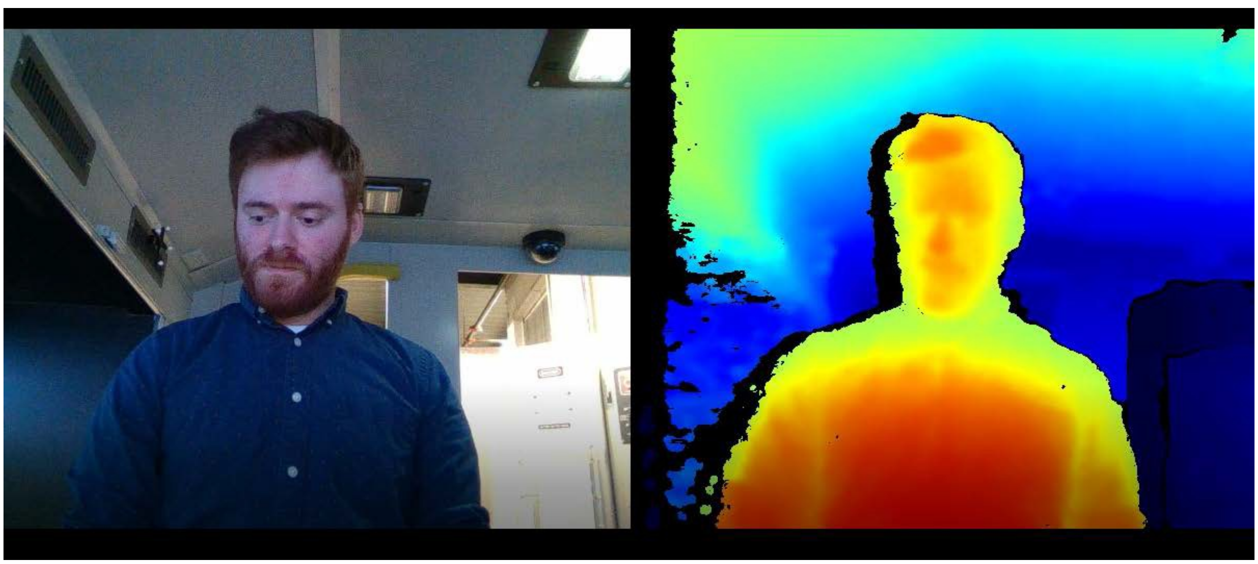


Figure 10: Video from Center Console – Standing Position

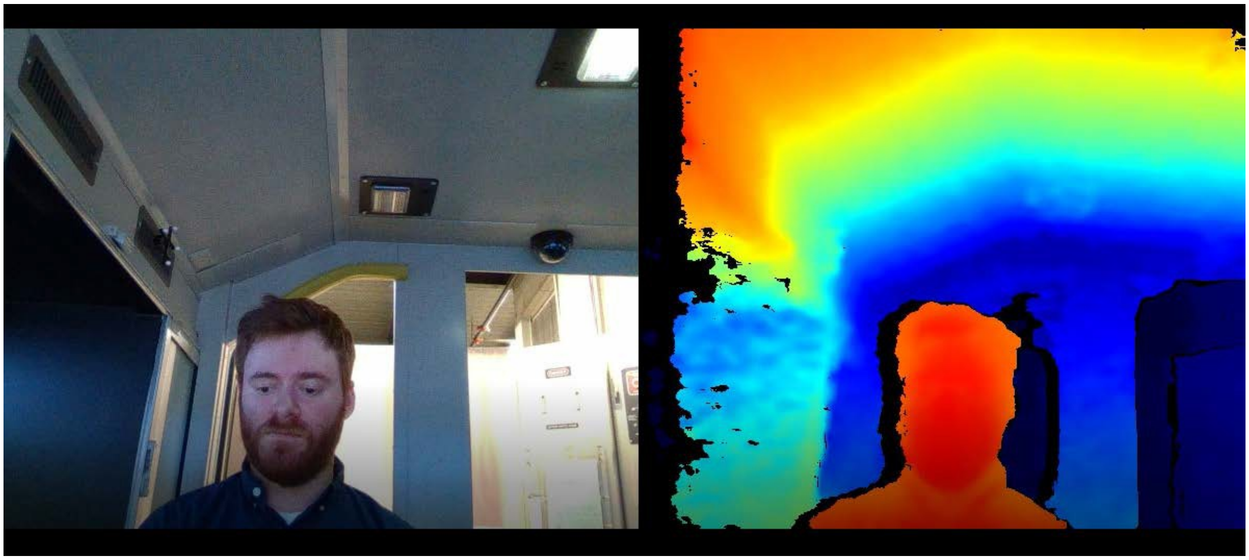


Figure 11: Video from Center Console - Sitting Position

These videos were run through basic facial feature tracking software to verify that the quality of video captured was sufficient for the purposes of this program. Additionally, the eye-tracking software was implemented on the videos that were acquired. Figure 12 shows a screen capture from this implementation, also showing blink capture and yawn ratio, both of which would be used to assist in the fatigue classification.

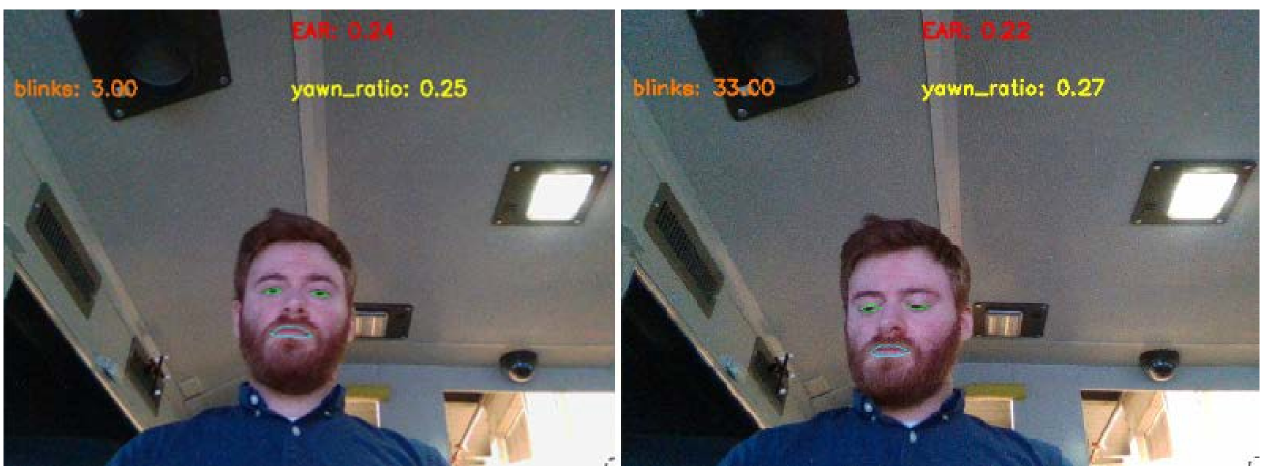


Figure 12: Video with eye-tracking software implemented

3.2 Reason for Discontinuation

The original MEFA proposal significantly relies on the use of people acting fatigued to generate sufficient data to train the train engineer fatigue classification algorithms. The research team had believed that this approach would be an efficient method for generating enough data to training the algorithm on common fatigue postures and behaviors. This method would also mitigate any risks posed to participants and researchers using traditional techniques to encourage participant drowsiness (e.g., working the "third shift," reducing nightly rest, etc.).

During Phase I the team conducted a literature review of different fatigue studies to evaluate the validity of this proposed approach. The papers reviewed covered several domains, including automobiles, trains, physiology research, and technology development (e.g. human factors, Ph.D. theses, the FRA eLibrary, signal and image processing, transportation conferences, etc.). The majority of works in the existing literature used highly ecologically valid conditions to measure drowsiness, such as working in simulators at night[20] or encouraging participants to come in a tired state.[21] Several studies used real-world data collection, relying on driver self-correction or drowsiness rather than full fatigue, so as to limit risk to participants.[22] One study used actors, but the focus of the study tested the effectiveness of the camera technology rather than the accuracy of the fatigue detection.[23]

Based on the evidence in the literature, researchers concluded that relying on the use of actors would increase the risk of developing a brittle algorithm that could not be reliably deployed in the field with confidence in its accuracy. Future studies must be designed in a way to sufficiently fatigue participants such that they are demonstrating ecologically valid physiological (e.g., eye closures, droopy head or body posture) and operationally erroneous behavioral cues (e.g., accelerating instead of decelerating).[24]

---

[20] https://dspace.mit.edu/handle/1721.1/27039

[21] http://www.robesafe.es/personal/bergasa/papers/IV201_Ivan.pdf

[22] https://ruor.uottawa.ca/bitstream/10393/23295/1/Abtahi_Shabnam_2012_thesis.pdf

https://core.ac.uk/download/pdf/82715006.pdf

http://www.iss.uni-stuttgart.de/forschung/veroeffentlichungen/friedrichs_iv2010.pdf

[23] https://www.researchgate.net/profile/Sumalini_Vartak/publication/305937386_Multi-View_Point_Drowsiness_and_Fatigue_Detection/links/57a6576108ae3f4529337fa4/Multi-View-Point-Drowsiness- and-Fatigue-Detection.pdf

[24] As a rule of thumb, there should be at least five to ten data points per physiological cue identified in Phase I as to generate enough training data to begin developing the classification algorithm. http://fastml.com/how- much-data-is-enough/

4. Conclusion

An engineer-monitoring system was developed by selecting a motion-capture device, identifying the location of fixation in the cabin, and evaluating the data with eye- and facial-tracking software. A logical step-by-step code was written based upon a procedure found within the GCOR. The methods by which data could be extracted from the CTIL Data Hub were identified and explored. These, in combination with Aurora's pre-existing procedure management module, would have allowed for a demonstration of the technology. However, it was determined midway through the program that engineers feigning fatigue would not have been suitable for creating a robust fatigue clarification algorithm.

A list of contributions from this program follows. Each of these will benefit future work in this area:

- In-depth review and report on current state of human-tracking devices
- Contextualized motion-capture strategies, including recommendations for specific device selection, mounting positions in the train cabin, and communication approaches
- Successful facial-tracking with the data from the motion-capture devices
- A completed DTR file of a GCOR activity
- Understanding of how to extract information about the cabin state from the CTIL Data Hub

For future work directly related to engineer fatigue, it is recommended that the scope include the analysis of engineers whom are actively fatigued. Based on the research from this program, relying on the use of actors would increase the risk of developing a brittle algorithm that could not be reliably deployed in the field with confidence in its accuracy. Additionally, the ability to acquire more detailed data from a source such as a TO or PTC, in addition to the data hub that was accessed during this program, would allow a more robust system by allowing for the construction of more engineer procedures to analyze through the locomotive position/information state machine.

5. References

[1] Whitlock, A. (2002, August 30). Driver Vigilance Devices: Systems Review. Railway Safety and Standards Board, Final Report by Quintec 03 T024 QUIN 22 RPT.

[2] P. W. Kithil, P.W. (2001, August 14-17). Driver Alertness Detection Research Using Capacitive Sensor Array. First Intl. Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

[3] Dawson, D., Searle, A.K., and Paterson, J.L. (2014). Look before you (s)leep: Evaluating the use of fatigue detection technologies within a fatigue risk management system for the road transport industry. *Sleep Medicine Reviews,18*(2), 141-152.

[4] Wang, Q., Yang, Y., Ren, M., and Zheng, Y. (2006). Driver Fatigue Detection: A Survey. 6th World Congress on Intelligent Control and Automation, Dalian, China.

[5] Barr, L., Popkin, S., and Howarth, H. (2009, June). An Evaluation of Emerging Driver Fatigue Detection Measures. Washington, DC: U.S. DOT/Federal Motor Carrier Safety Administration.

[6] Sikander, G., and Anwar, S. (2018). Driver Fatigue Detection Systems: A Review. *IEEE Transactions on Intelligent Transportation Systems,* 1-14.

[7] Ueno, H., Kaneda, M., and Tsukino, M (1994). Development of drowsiness detection system. *Proceedings of VNIS'94 – 1994 Vehicle Navigation and Information Systems Conference*, 15-20.

[8] Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., and Kirn, C.L. ("1994). Research on vehicle based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness. National Highway Traffic Safety Administration, Final Report DOT HS 808 247.

[9] Federal Highway Administration. (1998). FHWA-MCRT-98-006, PERCLOS: A Valid Psychophysiological Measure of Alertness As Assessed by Psychomotor Vigilance.

[10] Aboukhalil, A., Oman, C.M., Popkin, S., Pollard, J.K., and Howarth, H. (2005). Quantitative assessment of locomotive cab environment and engineer head movement for development of an alertness monitor. International Conference on Fatigue Management in Transportation Operations, Seattle, WA.

[11] Hanowski, R.J., Bowman, D., Alden, A., Wierwille, W.W., and Carroll, R. (2008). PERCLOS+: Development of a Robust Field Measure of Driver Drowsiness. 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting, New York.

[12] Schleicher, R., Galley, N., Briest, S., and Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired?" *Ergonomics, 51*(7), 982-1010.

[13] Tijerina, L., Gleckler, M., Stoltzfus, D., Johnston, S., Goodman, M.J., and Wierwille, W.W. (1999). A Preliminary Assessment of Algorithms for Drowsy and Inattentive Driver Detection on the Road. Washington, DC: National Highway Traffic Safety Administration.

[14] Johns, M.W., Tucker, A., Chapman, R., Crowley, K., and Michael, N. (2007). Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. *Somnologie, 11,* 234-242, 2007.

[15] S. Boverie, S., and A. Giralt, A. (2008). Driver vigilance diagnostic based on eyelid movement observation. *IFAC Proceedings Volumes, 41*(2), 12831-12836.

[16] Catalbas, M.C., Cegovnik, T., Sodnik, J., and Gulten, A. (2017). Driver fatigue detection based on saccadic eye movements. *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, pp. 913-917.

[17] Zils, E., Sprenger, A., Heide, W., Born, J., and Gais, S. (2005, September). Differential effects of sleep deprivation on saccadic eye movements" (in eng.). *SLEEP, 28*(9), 1109-15.

[18] Di Stasi, L.L., Renner, R., Catena, A., Cañas, J.J., Velichkovsky, B.M., and Pannasch, S. (2012). Towards a driver fatigue test based on the saccadic main sequence: A partial validation by subjective report data. *Transportation Research Part C: Emerging Technologies, 21*(1), 122-133.

[19] Xu, J. Min, J., and Hu, J. (2018). Real-time eye tracking for the assessment of driver fatigue. *Healthcare Technology Letters, 5*(2),54-58.

[20] Deng, L., Xiong, X., Zhou, J., Gan, P., and Deng, S. (2010). Fatigue Detection Based on Infrared Video Pupillography. *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, 1-4.

[21] Schumann, A., Ebel, J., and Bär, K.-J. (2017). Forecasting transient sleep episodes by pupil size variability. *Current Directions in Biomedical Engineering, 3*(2).

[22] Mora, G.G., Kortelainen, J.M., Hernández, E.R.P., Tenhunen, M., Bianchi, A.M., and Méndez, M.O. (2015). Evaluation of Pressure Bed Sensor for Automatic SAHS Screening. *IEEE Transactions on Instrumentation and Measurement, 64*(7), 1935-1943.

[23] Nakane, H., Toyama, J., and Kudo, M. (2011). Fatigue detection using a pressure sensor chair. *2011 IEEE International Conference on Granular Computing*, 490-495.

[24] Chen, R. (2013). *Sitting Behaviour-Based Pattern Recognition for Predicting Driver Fatigue* (Doctoral dissertation). Deakin University.

[25] Zhang, Y., and Hua, C. (2015). Driver fatigue recognition based on facial expression analysis using local binary patterns. *Optik, 126*(23), 4501-4505.

[26] Zhang, W., Murphey, Y.L., Wang, T., and Xu, Q. (2015). Driver yawning detection based on deep convolutional neural learning and robust nose tracking. *2015 International Joint Conference on Neural Networks,* 1-8.

[27] Fan, X., Sun, Y., Yin, B., and Guo, X. (2010). Gabor-based dynamic representation for human fatigue monitoring in facial image sequences. *Pattern Recognition Letters, 31*(3).

[28] Fan, X., Yin, B., and Sun, Y. (2007). Yawning Detection for Monitoring Driver Fatigue. *2007 International Conference on Machine Learning and Cybernetics,* 664-668.

[29] Solaz, J. et al., (2016). Drowsiness Detection Based on the Analysis of Breathing Rate

Obtained from Real-time Image Recognition. *Transportation Research Procedia, 14,* 3867-3876.

[30] Wusk, G., and Gabler, H. (2018). Non-Invasive Detection of Respiration and Heart Rate with a Vehicle Seat Sensor. *Sensors, 18*(5).

[31] Yu, X. (2012). Non-Contact Driver Drowsiness Detection System," Transportation Research Board, Final Report – Safety IDEA Project.

[32] Rogado, E., Garcia, J.L., Barea, R., Bergasa, L.M., and Lopez, E. (2009). Driver fatigue detection system. *2008 IEEE International Conference on Robotics and Biomimetics,* 1105-1110.

[33] Fu, R., and Wang, H. (2013). Detection of Driving Fatigue by Using Noncontact EMG and ECG Signals Measurement System. *International Journal of Neural Systems, 24*(3), 1450006.

[34] Dorrian, J., Lamond, N., Kozuchowski, K., and Dawson, D. (2008). The driver vigilance telemetric control system (DVTCS): Investigating sensitivity to experimentally induced sleep loss and fatigue. *Behavior Research Methods, 40*(4), 1016-1025.

[35] Lee, B.-L., Lee, D.-S., and Lee, B.-G. (2015). Mobile-based wearable-type of driver fatigue detection by GSR and EMG. *TENCON 2015 – 2015 IEEE Region 10 Conference,* Macao, China, 1-4.

[36] Khushaba, R.N., Kodagoda, S., Lal, S., and Dissanayake, G. (2011). Driver Drowsiness Classification Using Fuzzy Wavelet-Packet-Based Feature-Extraction Algorithm. *IEEE Transactions on Biomedical Engineering, 58*(1), 121-131.

[37] Wierwille, W.W., and Ellsworth, L.A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis & Prevention, 26*(5), 571-581.

[38] Dorrian, J., Roach, G.G., Fletcher, A., and Dawson, D. (2006). The effects of fatigue on train handling during speed restrictions. *Transportation Research Part F, 2006,* 243-257.

[39] Ji, Q., Lan, P., and Looney, C. (2006). A Probabilistic Framework for Modeling and Real-Time Monitoring Human Fatigue. *IEEE Transactions on Systems, Man, and Cybernetics, Pt. A, 36*(5), 862-875.

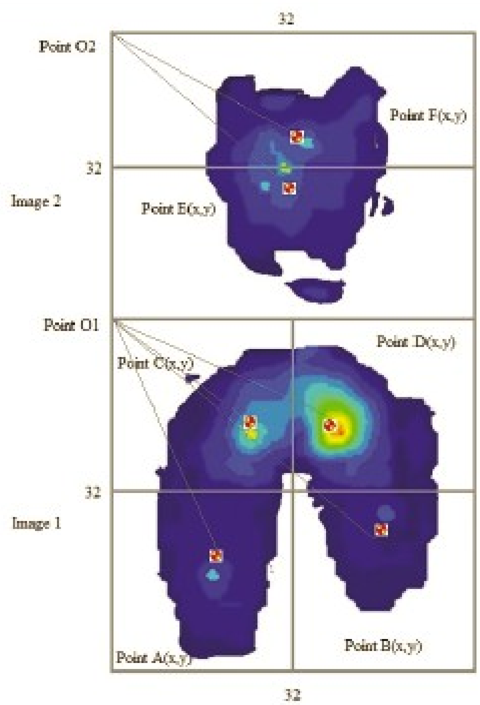
[40] Li, Z., Li, S.E., Li, R., Checng, B., and Shi, J. (2017). Online Detection of Driver Fatigue Using Steering Wheel Angles for Real Driving Conditions. *Sensors, 17*(3), 12.

[41] McDonald, A.D., Lee, J.D., Schwarz, C., and Brown, L.T. (2018). A contextual and temporal algorithm for driver drowsiness detection," *Accident Analysis & Prevention, 113,* 25-37.

Appendix A.
Pressure Sensor Images

List of features from the seat and backrest pressure distribution and acceleration data from the shoulder belt used by Chen [24] in a machine learning method to classify automobile driver fatigue.

Features from seat pressure sensor images from Chen [24]

Features (18-dimension)	Calculation
Load of image 1 (seat)	
Area of image 1	
Load of image 2 (backrest)	
Area of image 2	
Load of part A, B, C, D	
Distance from O1 to A, B, C, D	
Load of part E, F	
Distance from O2 to E, F	
Area of quadrilateral ABCD	
Distance from E to F	

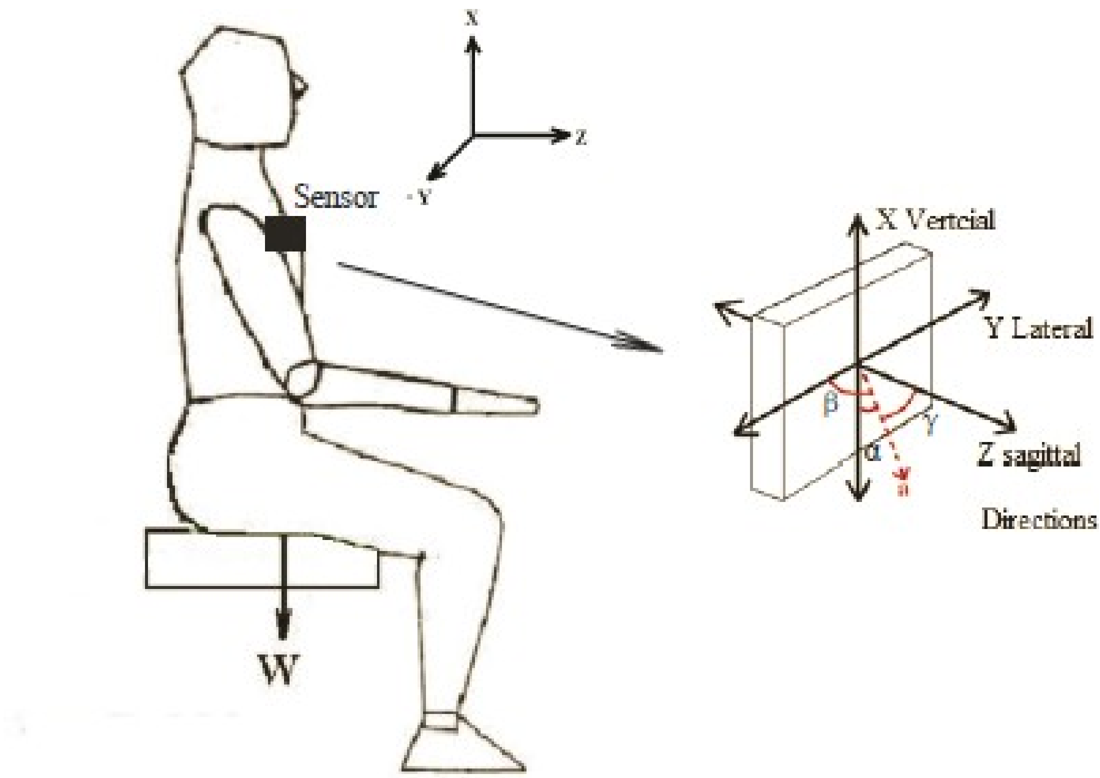
Note that there are four features each in Load of parts A, B, C, D, Distances from O1 for A, B, C, D, and two for Load of parts E,F and Distances from O2 to E and F.

Features derived from seatbelt-mounted accelerometer (3 signals)

(1) Vertical angle (angle about the X-axis)

(2) Lateral angle (angle about the Y-axis)

(3) Sagittal angle (angle about the Z-axis



Appendix B.
List of CTIL Variables

Variable	*Description*	*Units*	*Control Sta*
LOGTIME	Simulation time (From run start)	seconds	N
LOGDIST	Simulation distance (from run start)	km	N
LOGGRADE	The grade under the leading loco	m/m	N
LOGSPEED	Train speed	m/s	N
LOGACCEL	Train acceleration	m/s2 (meters per second squared)	N
LOGBCYL_IND	Brake cylinder from independent brakes	kPa	N
LOGBCYL_AUT	Brake cylinder from automatic brake	kPa	N
LOGBP	Brake pipe pressure at locomotive	kPa	N
LOGBCYL	Brake cylinder from independent and auto braking	kPa	N
TDSDIRN	Direction controller (Reverser)	(-1) Rev, 1 Fwd	Y
TDSNOTCH	Locomotive Notch Position	0-8 (0 being idle)	Y
CAN_HORN	1 - Horn on		Y
CAB_AUTO_MISM	1 - Release, 2 - min, 3-8 - service (continuous), 9 - full service, 10 - suppression, 17 - HO, 18 - Emergency	single	Y
LOGCURRENT	Traction motor current	Amps	N
ALM_ALERTER	1 - Alerter warning (Audio)	[binary]	N
CAN_ALERT_RESET	1 - Alerter pressed	[binary]	Y
ALERT_PENALTY	1 - Penalty signaled by alerter system	[binary]	N
DSR_COM005	Value changes when video recorder turns on.	n/a	N
ALM_BELL	1 - Bell on	[binary]	Y
LOGFLOW	Air flow	kg/sec	N
CAN_BAIL	1 - Bail pressed	[binary]	Y
IND_HAND_PERCENT	Percentage of independent applied (0.0 - 1.0)	percentage	Y
LOGROUT_DRAFTMAX	Max run-out draft force in train in log interval	Newtons	N
LOGROUT_MAXPOS	Position in train of maximum run-out force	(Car number, integer)	N
LOGRIN_BUFFMAX	Max Run-in buff force in train in log interval	Newtons	N
LOGRIN_MINPOS	Position in train of maximum run-in force	(Car number, integer)	N
LOGSS_DRAFTMAX	Max stead state draft force in train in log interval	Newtons	N
LOGSS_BUFFMAX	Max stead state buff force in train in log interval	Newtons	N
LOGFUEL,	Total fuel consumed by all locomotives	Litres	N
LOGFUEL_RATE	Fuel consumption for all locos-instantaneous rate	Litres/min	N
GETO_TOT_TIME	Total time TO has been running	seconds	N
GETO_TIME_AUTO	Total time TO has been running in automatic mode	seconds	N
GETO_AUTO_AV	TO: value unknown to us.	seconds	N
GETO_TIME_MAN_ONLY	Total time TO has been running in manual mode	seconds	N

Appendix C.
Yellow Flag Procedure DTR File


```xml
<?xml version="1.0" encoding="UTF-8"?>
<alias name="Train">
    <subsystems>
        <perception>
            <enumList/>
            <primitives>
                <primitive id="2002" name="Speed">
                    <units max="250.0" min="-250.0" name="Meter_sec" type="F4"/>
                </primitive>
                <primitive id="2003" name="Acceleration">
                    <units max="90.0" min="-90.0" name="Meter_sec_2" type="F4"/>
                </primitive>
            </primitives>
        </perception>
        <actuation>
            <enumList/>
            <primitives/>
        </actuation>
        <core>
            <enumList/>
            <primitives/>
        </core>
        <hmi>
            <enumList/>
            <primitives/>
        </hmi>
    </subsystems>
    <common>
        <constants/>
    </common>
    <transitions startState="300">
        <transition current="300" next="301">
            <condition>Default</condition>
        </transition>
        <transition current="301" next="302">
            <condition>Default</condition>
        </transition>
    </transitions>
    <route>Android/data</route>
    <procedures>
        <process id="0" isContingency="false" name="Flight Procedure">
            <roleset id="0">
                <automation defaultLevel="0"/>
                <priority defaultLevel="0"/>
                <operations>
                    <!--Subprocedure: Start Operation-->
                    <operation id="1" type="3">
                        <parentSet/>
                        <procedureId result="1">300</procedureId>
                        <else result="2"/>
                    </operation>
                    <!--Subprocedure: Nominal Operation-->
                    <operation id="2" type="3">
                        <parentSet>
                            <parentCondition opId="1" opResult="1"/>
                        </parentSet>
                        <procedureId result="1">301</procedureId>
                        <else result="2"/>
                    </operation>
```

```
            <!--Subprocedure: Stop Operation-->
            <operation id="3" type="3">
                <parentSet>
                    <parentCondition opId="2" opResult="1"/>
                </parentSet>
                <procedureId result="1">302</procedureId>
                <else result="2"/>
            </operation>
        </operations>
    </roleset>
    <roleset id="1">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations>
            <!--Subprocedure: Start Operation-->
            <operation id="4" type="3">
                <parentSet/>
                <procedureId result="1">300</procedureId>
                <else result="2"/>
            </operation>
            <!--Subprocedure: Nominal Operation-->
            <operation id="5" type="3">
                <parentSet>
                    <parentCondition opId="4" opResult="1"/>
                </parentSet>
                <procedureId result="1">301</procedureId>
                <else result="2"/>
            </operation>
            <!--Subprocedure: Stop Operation-->
            <operation id="6" type="3">
                <parentSet>
                    <parentCondition opId="5" opResult="1"/>
                </parentSet>
                <procedureId result="1">302</procedureId>
                <else result="2"/>
            </operation>
        </operations>
    </roleset>
</process>
<process id="300" isContingency="false" name="Start Operation">
    <roleset id="0">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
    <roleset id="1">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
</process>
<process id="301" isContingency="false" name="Nominal Operation">
    <roleset id="0">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
    <roleset id="1">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
</process>
```

```xml
<process id="302" isContingency="false" name="Stop Operation">
    <roleset id="0">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
    <roleset id="1">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
</process>
<process id="400" isContingency="true" name="Yellow Flag">
    <roleset id="0">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations>
            <!--WAIT_FOR  Speed  <  10-->
            <operation actor="1" display="true" id="7"
                monitor="2" time="30" type="2">
                <name>WAIT_FOR  Speed  &lt;  10</name>
                <parentSet/>
                <id>2002</id>
                <value result="1">DID2002 &lt; 4.4704</value>
                <else result="2"/>
            </operation>
            <!--[Custom]  WAIT_FOR  Green Flag  =  Passed-->
            <operation display="true" id="8" monitor="1"
                time="30" type="2">
                <name>[Custom]  WAIT_FOR  Green Flag  =  Passed</name>
                <parentSet>
                    <parentCondition opId="7" opResult="2"/>
                </parentSet>
                <value>DID500 == 5</value>
                <output type="2">
                    <id>Green Flag</id>
                    <value result="1">Green Flag = Passed</value>
                </output>
                <else result="2"/>
            </operation>
            <!--[Custom]  WAIT_FOR  Distance  >=  4 miles-->
            <operation display="true" id="9" monitor="0"
                time="30" type="2">
                <name>[Custom]  WAIT_FOR  Distance  &gt;=  4 miles</name>
                <parentSet>
                    <parentCondition opId="7" opResult="2"/>
                </parentSet>
                <value>DID500 == 5</value>
                <output type="2">
                    <id>Distance</id>
                    <value result="1">Distance &gt;= 4 miles</value>
                </output>
                <else result="2"/>
            </operation>
            <!--[Custom]  WAIT_FOR  Dispatch  ==  Verified Safe-->
            <operation display="true" id="10" monitor="1"
                time="30" type="2">
                <name>[Custom]  WAIT_FOR  Dispatch  ==  Verified Safe</name>
                <parentSet>
                    <parentCondition opId="9" opResult="2"/>
                </parentSet>
                <value>DID500 == 5</value>
                <output type="2">
```

```
                    <id>Dispatch</id>
                    <value result="1">Dispatch == Verified Safe</value>
                </output>
                <else result="2"/>
            </operation>
            <!--WAIT_FOR  Speed  >  10-->
            <operation actor="1" display="true" id="11"
                monitor="2" time="30" type="2">
                <name>WAIT_FOR  Speed  &gt;  10</name>
                <parentSet>
                    <parentCondition opId="8" opResult="2"/>
                    <parentCondition opId="10" opResult="2"/>
                </parentSet>
                <id>2002</id>
                <value result="1">DID2002 &gt; 4.4704</value>
                <else result="2"/>
            </operation>
        </operations>
    </roleset>
    <roleset id="1">
        <automation defaultLevel="0"/>
        <priority defaultLevel="0"/>
        <operations/>
    </roleset>
</process>
</procedures>
<contingencies>
    <!--Yellow Flag-->
    <contingency id="401" name="Yellow Flag" procedureId="400">
        <operations/>
    </contingency>
</contingencies>
</alias>
```

Abbreviations and Acronyms

ALIAS	Aircrew Labor In-cockpit Automation System
CTIL	Cab Technology Integration Lab
DTR	Digital Type-Rating
FOV	Field Of View
FRA	Federal Railroad Administration
MEFA	Monitoring Engineer Fatigue
MIT	Massachusetts Institute of Technology
MPH	Miles Per Hour
PTC	Positive Train Control
TO	Trip Optimizer