https://doi.org/10.2352/EI.2024.36.17.AVM-112

Exploring Object Detection and Image Classification Tasks for Niche Use Case in Naturalistic Driving Studies

Ryan Peruski, Deniz Aykac, Lauren Torkelson, Thomas Karnowski
Oak Ridge National Laboratory, Oak Ridge, TN 37831

Abstract

Naturalistic driving studies consist of drivers using their personal vehicles and provide valuable real-world data, but privacy issues must be handled very carefully. Drivers sign a consent form when they elect to participate, but passengers do not for a variety of practical reasons. However, their privacy must still be protected. One large study includes a blurred image of the entire cabin which allows reviewers to find passengers in the vehicle; this protects the privacy but still allows a means of answering questions regarding the impact of passengers on driver behavior. A method for automatically counting the passengers would have scientific value for transportation researchers. We investigated different image analysis methods for automatically locating and counting the non-drivers including simple face detection and fine-tuned methods for image classification and a published object detection method. We also compared the image classification using convolutional neural network and vision transformer backbones. Our studies show the image classification method appears to work the best in terms of absolute performance, although we note the closed nature of our dataset and nature of the imagery makes the application somewhat niche and object detection methods also have advantages. We perform some analysis to support our conclusion.

Keywords: *Naturalistic driving studies, privacy, object detection, image classification*

** Corresponding author's E-mail: karnowskitp@ornl*.*

Introduction

Traffic fatalities are a major problem across with world, with 35,000 annual deaths in the United States of America alone [1]. Insight into how crashes and near-crashes occur can be obtained by collections of actual driving data in the field; such studies are called Naturalistic Driving Studies (NDS), with data collected including temporal data on vehicle dynamics, as well as radar and video data [2]. The Second Strategic Highway Research Project (SHRP2) was conducted with approximately 3000 drivers in 6 data collection sites in the US between 2010 and 2013 [3]. The size of the study makes it a valuable resource for fundamental questions about driver behavior, even with new advances in safety equipment.

One question involves the impact of passengers on driver behavior. Intuitively, passengers could act as distractions, but they could also assist in driving by acting as a "second set" of eyes to alert the drivers to issues. Reported research has found various factors; a study sampling car crashes found most crashes were due to inattentiveness, and a major component was when drivers were conversing with passengers [4]. Other studies have shown similar results, with one showing passenger distraction as the most common distraction by a large margin (7.4% in distraction from passenger vs. *2.2% in the next highest distraction: using a mobile phone) [9] and another showing distractions from passengers being observed in 53.2% of vehicles with passengers present [10]. There are many variables involved, including the ages of the passenger / driver and the nature of the driving itself, but the overall impact remains an open question.

The SHRP2 NDS included a blurred cabin image (Figure 1) sized 360x480 pixels, which was captured every 10 minutes to allow for the detection of passengers in the vehicle. Due to the nature of the consent agreements, passengers were not specifically included and therefore their privacy must be preserved. (Note that the image in Figure 1 is an example image and is not from the study itself). Thus, the blurred cabin image features a means of privacy protection, and further removing the blur was not an option due to the privacy constraints.

Figure 1*: An example of a blurred SHRP2 cabin image to show the technique. Note this is not an image of an actual NDS subject, but rather is an image used to illustrate the effect [3].*

This is an example of a niche "downstream task" for computer vision [7], which would have been more difficult prior to the advent of modern computer vision and machine learning tools. In this case we wish to determine the number of passengers both in the front seat and rear seats of the vehicles to help identify NDS "trips" which could help yield information about passenger impacts on behavior. An initial study was conducted which used an image classification approach using the AlexNet architecture [23] for transfer learning by partitioning the image into regions of interest with a dataset selected from example images in the SHRP2 study. Subsequent work [11] used an object detection approach, which has the advantages of helping to locate anomalies in passenger positions or gaining better understanding of other potential safety issues such as unusual seating arrangements. Overall, we believe the object detection approach has advantages over an image classification approach, but it also has some potential issues including the difficulty of detecting the passengers, particularly in the rear of the vehicle. Thus, in this work, we sought to determine how the object detection approach compared with robust image classification methods, especially given that the overall environment is fairly constrained. We were interested in comparisons with newer models used as backbones [5] [12][13] and combined language-imagery methods [8], with the latter offering promise for zero-shot applications where essentially no training or fine-tuning is needed. While there is considerable volume of data in SHRP2, finding and labeling data for training sets is always difficult and costly, so methods which use less data for training have advantages, especially in niche applications. Another issue is the "unnatural nature" of the imagery, as the self-imposed blur is not typical of large image datasets used in training large models.

The paper is organized as follows. First, we discuss the datasets and dataset preparation. We next review the methods used in the evaluation. We follow with a summary of the results and discussion of the method performances. Finally, we conclude with additional overall discussion and suggestions for follow-on work.

Data

Dataset

The SHRP2 data collection system used a set of analog cameras to capture video at 15 frames per second. A single blurred snapshot of the entire cabin was taken every 10 minutes to provide situational awareness of the cabin environment, particularly the presence and location of passengers. The entire SHRP2 dataset is archived at Virginia Tech Transportation Institute (VTTI) [25].

The dataset used was a collection of day and night images from most of the vehicles in the SHRP2 dataset (some vehicles were omitted due to issues such as a faulty camera or hang tags which blocked the view or had identifying information). The data was collected from the VTTI archive in three phases. The first phase was a collection where there were known to be passengers in the vehicle. The second was a collection of a daytime and nighttime image from every vehicle in the dataset. These two data collections consisted of 2834 different vehicles with a median of 4 images per vehicle [24]. The third was a final collection obtained from the archived data in 2023 which also used day and night views from each vehicle but were checked to ensure they had not been used in any earlier datasets, so we were certain they were images that could be used for testing purposes. This data collection consisted of 1081 images from 167 different vehicles with a median of 6 images per vehicle. Although the images are new, all 167 vehicles were present in the other two data collections. We used the first two sets for training and validation and the final set for testing.

There were sometimes small changes in the camera position, but generally we were confident that the front passenger could be assessed with a region of interest covering the left half of the image. The front driver was in the right half, but the driver was always present, so we did not really need to estimate their presence. However, it served as a good data source for "person present" in the training set for the front-seat passengers.

The rear vehicle locations were more difficult (and partially a reason why the object detection approaches could have benefits). An example of a single vehicle with and without backseat passengers is shown in Figure 2. There were typically 3 positions to assess, but some vehicles (large vans) had as many as six, and some vehicles (trucks with no rear seat) had no rear passengers. Further, the backseat data was unbalanced (prevalence < 0.10). We did some initial experiments by balancing the data. We found that models performed better with the larger unbalanced datasets, so we used the original unbalanced data. A summary of the datasets is shown in Table 1, with the number of images, unique vehicles, and occurrences of front passengers and rear passengers listed.

Table 1. Summary of Datasets

Name	Images	Vehicles	Front	Rear
Witcher	5654	2834	1515	666
2023	1081	167	375	221

Establishing Ground Truth

We established ground truth for passenger presence in two phases. For the front seat, we split the images in half and manually checked presence or absence of people (we checked the driver side as well in the process). For the backseat, locations were selected for each vehicle to establish the likely positions of passengers. This was a difficult procedure because in many cases it was not clear when passengers were present, because the expected signs of occupancy were not consistent; for example, often a rear passenger face was hidden due to occlusion from the front seat. We examined the images from identical vehicles for comparison purposes which helped us detect the differences in the back seat occupancy when applicable. This per-vehicle location was useful for the image classification approaches, with a 91x91 pixel region of interest selected. For the object detection ground truth, the process was slightly different. We selected the center of the face for the front seat passenger and a 275x320 pixel region of interest. Finally, we note that there were still cases where the 'ground truth' was not clear. In these cases, we omitted this area of interest (but not necessarily the entire image).

Methods

Object Detection and Location

We tested two object detection methods which targeted face or body detection. First, we used a recently developed model from researchers at VTTI [11]. The method was specifically fine-tuned for the SHRP2 blurred cabin imagery and used a Faster-RCNN

model [14], with a ResNet50 backbone (23M parameters). We used the tool implementation from GitHub [21] and simply processed our newest dataset for comparison with other methods. The second method was an "out of the box" face detection tool "RetinaFace" [15][16] as we had found it operated rather robustly against other NDS data [22]. RetinaFace was used "as is", with no fine-tuning, as a baseline method; the implementation also used a ResNet50 backbone.

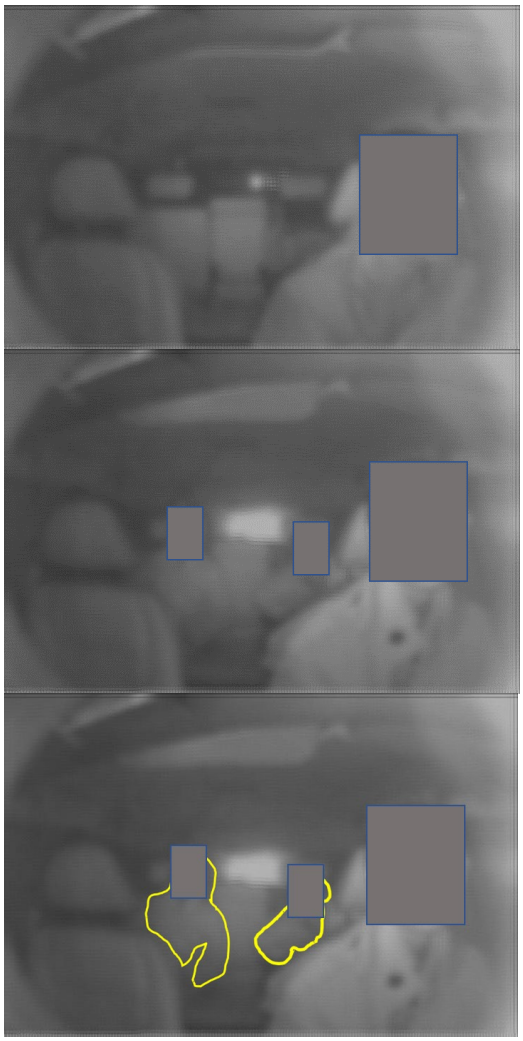

Figure 2. Image example from study. *Top: Image from a vehicle with no backseat passengers. The face of the front seat driver is occluded by a box added by the authors. Middle: Image from the same vehicle but with two back-seat passengers. Again, faces are occluded, but they are visible in the actual image, along with body components occluding parts of the back seat. Bottom: Same image as middle, but with the back seat passengers outlined to accentuate their presence.*

Image Classification

We used three general approaches for image classification backbones: convolutional neural network (CNN), Vision Transformer (ViT), and contrastive language-image pretrained (CLIP) transformer models. In our earlier unpublished work, which was the basis for comparison in [11], we used AlexNet for transfer learning, but newer models are available that should be more effective. We performed some initial testing to determine which models could be easily implemented and fine-tuned for this problem. We selected pre-trained ResNet18 (to represent a smaller sized network) [12][20] and EfficientNetB7 (a newer, larger network) [13][19]. ResNet, or Residual Network, provides shortcuts through residual learning, allowing for less complexity and easier optimization. ResNet18, a form of ResNet with 18 layers, represents an earlier topology with advances over other models such as AlexNet [23], and EfficientNetB7 features a newer topology with lower computing power and parameter requirements than comparably performant models. EfficientNetB7 was shown to outperform ResNet50 on the ImageNet dataset. However, ResNet18, the scaled-down version of the ResNet50, is a good baseline for transfer learning, as it does not take long to train. As a rough measure of the magnitude of these models, ResNet18 has 11.4M parameters and EfficientNetB7 has 66M parameters. We also used vision transformers (ViTs) due to their recent popularity in downstream tasks. ViTs use the transformer architecture which is simpler than CNN architectures [5][6] but more scalable to larger datasets to constitute "foundation" models [7]. Our implementation uses a ViT B-16 model [5] with 86M parameters. We used a PyTorch implementation to fine-tune and test these models.

For fine tuning these image classification methods, we created a training set of 85% of the Witcher data and reserved the rest for a validation set. We used a consistent batch size of 4, max epoch amount to 1000, and a patience of 10, based on validation accuracy. We also used random flips, affine, and color jitters for data augmentation, as well as a resize, center crop, and normalization for image transforming. We used a learning rate of 0.01 for all methods. In all cases, ten fine-tuning trials were run, and the best performing model based on the validation score was chosen.

Finally, we used the Contrastive Language and Image Pretraining (CLIP) [8][18] model with a ViT backbone (totaling approximately 63M parameters). CLIP embeds images and text information together when training. We used the Vision Transformer B16 model [17] for the visual component and distilBERT-base-uncased [27][26] for the language component. We first tested some preliminary prompts (see Table 2), ultimately selecting "a picture of a person" and "a picture of an empty seat". Then, we fine-tuned the model with our own training data, and then tested it again analogous to the other image classifier methods. We normalized the scores for these metrics based on the training data results to ensure scores ranged from 0 to 1 for the estimates. We used a learning rate of 0.00001 for the CLIP method.

The initial prompt tests were performed somewhat ad-hoc, using both our intuition and suggestions generated by ChatGPT4 [28]. Table 2 shows example prompts along with the F1 score. While the best prompt was "A picture of a blurred empty seat" and "A blurred picture of a person" with an F1-score of 0.67, we elected to use the non-blur version of the prompts as the difference was minimal and we wanted to do more of a comparison leveraging natural images.

Evaluation Comparisons

Comparing the broad image classification and object detection methods required some attention. The detections of the object detectors were evaluated by comparing the detected and ground truth bounding boxes with an Intersection over Union score. Each object detection was tested against the ground truth and the highest IOU was utilized. The confidence score of this best detection was logged for comparison with the image classification methods.

Table 2. Zero-Shot Prompt Testing

No Person Prompt	Person Prompt	F1
"Empty front passenger seat in a car"	"A car's front passenger seat occupied by a person"	0.40
"Car's front passenger seat without any person"	"Interior of a car with someone sitting in the front passenger seat"	0.49
"Automobile interior with no one in the front passenger seat"	"Car with a person in the passenger seat"	0.58
"Unoccupied front passenger seat of a vehicle"	"Automobile interior showing a person in the front right seat"	0.32
"Front seat of a car next to the driver is empty"	"An individual is sitting in the front seat of a car next to the driver"	0.55
"A picture of an empty seat"	"A picture of a person"	0.61
"A picture of a blurred empty seat"	"A blurred picture of a person"	0.67
"Blurred Automobile interior with no one in the front passenger seat"	"Car with a blurred person in the passenger seat"	0.56

Results and Discussion

Front Passengers

The precision-recall (PR) and receiver-operating characteristic (ROC) curves are shown in Figure 3 and Figure 4. Results are tabulated as well in Table 3, where we show the area under the curve (AUC) for the PR and ROC curve, the F1 score, and the precision and recall for the best F1 score. We also show FPR*, which is the false positive rate when the true positive rate is 0.78, for comparison with [11]. From these results, we see that both the fine-tuned image classification and object detection methods perform well, with the baseline RetinaFace performing well but missing many of the passengers. The CLIP zero-shot does have some promise but shows inferior performance; out of the two non-fine-tuned methods, the RetinaFace does seem to perform better, but it was specifically trained for face detection. Of the Image Classification methods, excluding CLIP Zero-Shot (AUC = 0.67), each method performed very similarly, although EfficientNetB7 did perform the best (AUC = 1.00). Of the Object Detection methods, there was a significant increase in performance from Retinaface to the Faster-RCNN Object Detector from VTTI (AUC = 0.83 vs. AUC = 0.75).

Rear Passengers

The rear passenger results are shown in Figure 7 and Figure 8 and tabulated in Table 4. The back seat task was understandably more difficult than the front seat due to the lower resolution, occlusion, and general poor visibility. We note the CLIP Zero-Shot method performs similar for the ROC AUC for the front passenger case (AUC=0.67), but the RetinaFace method essentially does not detect rear passengers. Overall, the image classification methods again perform better than the object detection methods. The ResNet18, CLIP Fine-Tuned, and ViT B 16 methods all performed similarly, and the best compared to other methods (AUC = 0.90, 0.89, 0.89, respectively), followed by EfficientNetB7 (AUC = 0.87). For Object Detection methods, both methods performed similarly to random guessing, although VTTI Faster-RCNN performed better

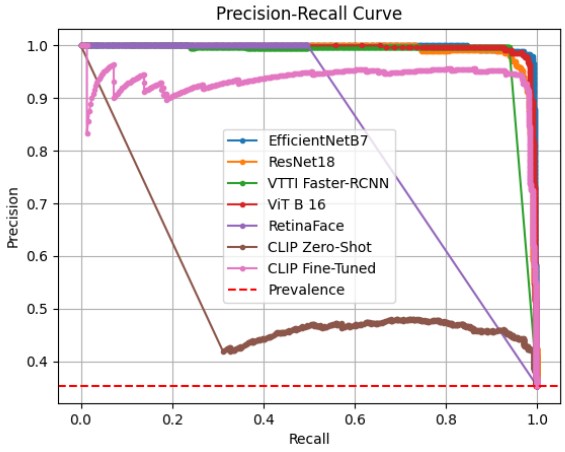

Figure 3 *Perf-Recall curves for each method – Front Seat¬*

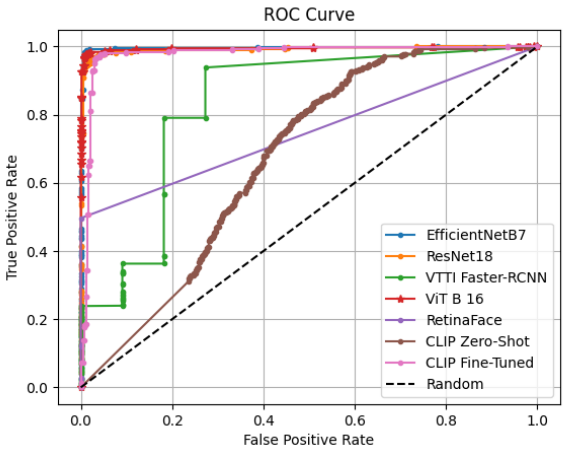

Figure 4 *ROC curves for each method – Front Seat*

Table 3. Front Seat Performance

Method	PR AUC	ROC AUC	F1	Prec	Rec	FPR*
VTTI Faster-RCNN	0.97	0.83	0.97	1.00	0.94	0.18
EfficientNet B7	1.00	1.00	0.98	0.98	0.99	0.00
Vision Transformer B 16	0.99	0.99	0.98	0.98	0.98	0.00
RetinaFace	0.77	0.75	0.66	1.00	0.50	0.56
CLIP Zero-Shot	0.54	0.67	0.61	0.46	0.93	0.46
CLIP Fine-Tuned	0.93	0.98	0.95	0.94	0.97	0.02
ResNet18	0.99	0.99	0.96	0.98	0.94	0.00

than the Retinaface (AUC = 0.56 vs. AUC = 0.51). Again, the backseat passenger prevalence level is very low (prevalence < 0.10), so guessing randomly gives around an accuracy of 90%. Therefore, the difference in accuracies for the backseat are not as drastic as the front seat.

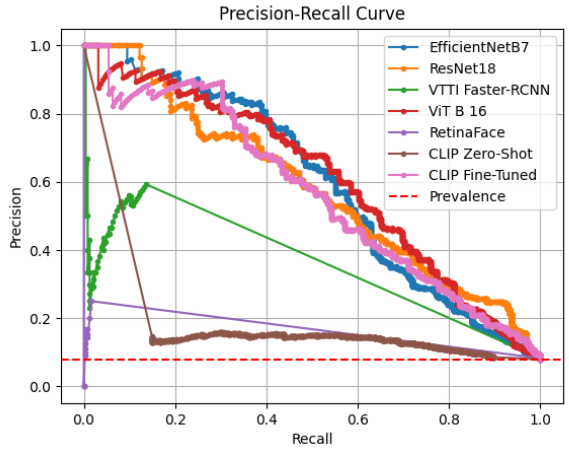

Figure 5 *Perf-Recall curves for each method – Back Seat*

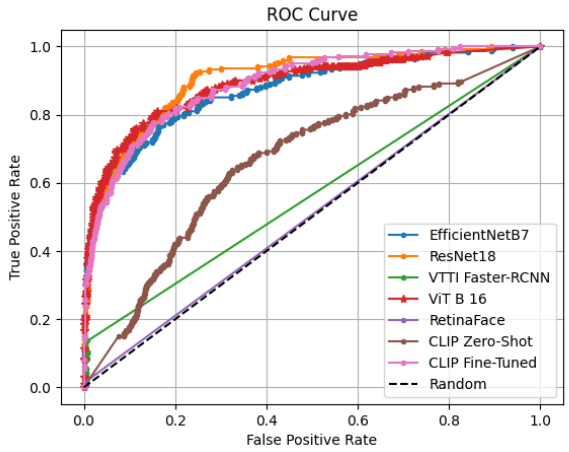

Figure 6 *ROC curves for each method – Back Seat*

Dataset Size Impact

As a final test, we were interested in determining the impact of different amounts of training data. In many cases in real-world problems, data is limited and labeled training data can be costly, so methods that require less data should be favored. In particular, we reasoned that larger models should require less training data since they have essentially "learned" more effectively in their pre-training stages. To test this, we ran 3 trials on lower percentages of training data by randomly removing 5%, 10%, 25%, 50%, and 75% of data from the Witcher set, and again picking the best performer in the validation set, then testing against the test sets (which were not reduced in size). Other procedures were identical.

Discussion

As expected, each method did experience a decrease in performance when going from the front passengers to the back passengers, likely due to obstructed views (behind the front seats), lower resolution, and possibly cases where the passenger was not within the ideal regions of interest for each vehicle.

Table 4. Back Seat Performance

Method	PR AUC	ROC AUC	F1	Prec	Rec	FPR*
VTTI Faster-RCNN	0.36	0.56	0.22	0.59	0.14	0.75
EfficientNet B7	0.59	0.87	0.58	0.61	0.55	0.19
Vision Transformer B 16	0.61	0.89	0.59	0.67	0.53	0.15
RetinaFace	0.16	0.51	0.15	0.08	1.00	0.78
CLIP Zero-Shot	0.20	0.67	0.24	0.15	0.56	0.54
CLIP Fine-Tuned	0.56	0.89	0.54	0.57	0.52	0.17
ResNet18	0.58	0.90	0.55	0.52	0.58	0.16

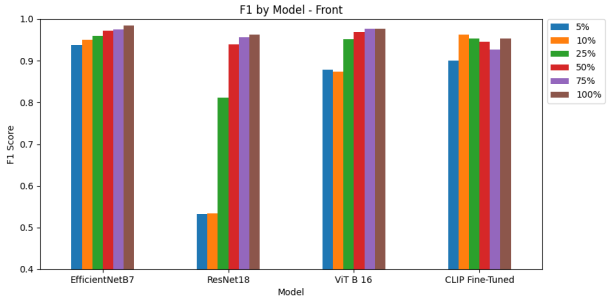

Figure 7 *F1 scores for each training data size – Front Seat*

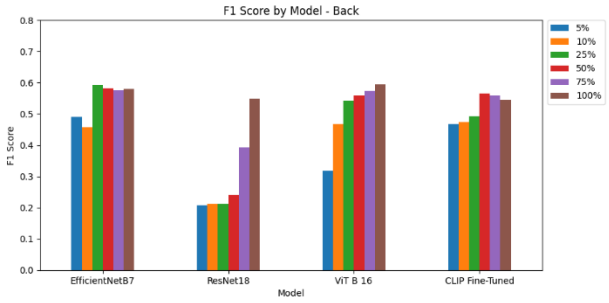

Figure 8 *F1 scores for each training data size – Back Seat*

The results are summarized in Figure 7 and Figure 8. (Note that the vertical scales are different on both these plots.) For the front passengers, the EfficientNetB7 and CLIP Fine-Tuned methods seem to be noticeably the most resilient to smaller training data sizes. The ResNet18 is the least resilient to the smaller data set sizes. For the rear passengers, a similar effect is shown with ResNet18 showing the least resilience to small training data sizes, while EfficientNetB7 and CLIP Fine-Tuned methods show the most resilience with the back passengers. We note that the CLIP zero-shot model had a ROC AUC very similar from the front passengers to the back (AUC for front and back = 0.67). However, its F1 score was significantly worse for the back seat case (F1 of 0.61 vs 0.24).

The four fine-tuned image classification methods performed both tasks similarly, but ResNet18 had issues with smaller training data. Finally, we note an object detector could be retrained with CLIP and VIT as well.

Conclusions

The image classification methods outperformed the object and face detection methods. However, the image classification methods, regardless of their performance, still have a distinct disadvantage in that they presuppose the locations of passengers in the vehicles. This is likely not a practical issue for a controlled or semi-controlled environment such as within vehicles, or in some industrial applications, but generally an object detection method would be preferred based on its ability to generalize to the location variability. A combination of the two may be the best method, however. We acknowledge that our dataset is closed and that the way we performed this study may not be an option for many application cases. The use of pre-existing detections such as RetinaFace or the CLIP zero-shot models did not seem to perform well, which is understandable given the "non-natural" nature of the imagery. Due to the nature of the images, no matter the method, fine-tuning seems to be imperative to getting effective results. Zero-shot functions may serve well for more "natural" images, or to get quick classifications.

Acknowledgments

We acknowledge the assistance of Virginia Tech Transportation Institute, and staff at FHWA Turner Fairbank. This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan). Work was funded by the Federal Highway Administration of the US Department of Transportation, Exploratory Advanced Research Fund.

References

[1] "Early Estimate of Motor Vehicle Traffic Fatalities for the First Half of 2021," https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813199

[2] Guo, Feng. "Statistical methods for naturalistic driving studies." Annual review of statistics and its application 6 (2019): 309-328.

[3] Hankey, Jonathan M., Miguel A. Perez, and Julie A. McClafferty. *Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets*. Virginia Tech Transportation Institute, 2016.

[4] Neyens, D. et al, "The influence of driver distraction on the severity of injuries sustained by teenage drivers and their passengers", Accident Analysis & Prevention, Volume 40, Issue 1,2008, Pages 254-259, ISSN 0001-4575, https://doi.org/10.1016/j.aap.2007.06.005.

[5] Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929 (2020).

[6] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).

[7] Gan, Zhe, et al. "Vision-language pre-training: Basics, recent advances, and future trends." Foundations and Trends in Computer Graphics and Vision 14.3–4 (2022): 163-352.

[8] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In International conference on machine learning (pp. 8748-8763). PMLR.

[9] Sullman, M. J. (2012). An observational study of driver distraction in England. Transportation research part F: traffic psychology and behaviour, 15(3), 272-278.

[10] Huisingh, C., Griffin, R., & McGwin Jr, G. (2015). The prevalence of distraction among passenger vehicle drivers: a roadside observational approach. Traffic injury prevention, 16(2), 140-146.

[11] Papakis, I. et al. (2021). Convolutional Neural Network-Based In-Vehicle Occupant Detection and Classification Method using Second Strategic Highway Research Program Cabin Images. Transportation Research Record, 2675(8), 443–457.

[12] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[13] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.

[14] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.

[15] Deng, J., Guo, J., Zhou, Y., Yu, J., Kotsia, I., & Zafeiriou, S. (2019). Retinaface: Single-stage dense face localisation in the wild. arXiv preprint arXiv:1905.00641.

[16] https://pypi.org/project/retina-face/

[17] https://pytorch.org/vision/main/models/generated/torchvision.models.vit_b_16.html#torchvision.models.vit_b_16

[18] Moein Shariatnia. (2022). moein-shariatnia/OpenAI-CLIP: openai-clip-first-release (v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.6845731

[19] https://pytorch.org/vision/main/models/efficientnet.html

[20] https://pytorch.org/vision/main/models/generated/torchvision.models.resnet18.html

[21] https://github.com/VTTI/object-detection

[22] Karnowski, Thomas, et al. Database to Enable Facial Analysis for Driving Studies (DEFADS). No. ORNL/TM-2022/2786. ORNL, Oak Ridge, TN (United States), 2022.

[23] Krizhevsky, A. et al. "ImageNet classification with deep convolutional neural networks." Communications of the ACM 60.6 (2017): 84-90.

[24] Witcher, Christina; Perez, Miguel A.; Karnowski, Thomas; Ferrell, Regina; Aykac, Deniz, 2020, "Blurred Cabin Imagery for Passenger Detection in SHRP2 NDS", https://doi.org/10.15787/VTT1/OONZ5I, VTTI, V2.

[25] https://insight.shrp2nds.us/

[26] https://huggingface.co/distilbert-base-uncased

[27] Sanh, Victor, et al. "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter." arXiv preprint arXiv:1910.01108 (2019).

[28] OpenAI. (2023). GPT-4. https://www.openai.com/models/gpt-4/

Author Biography

Ryan Peruski is working toward a BS and MS in Computer Science at the University of Tennessee Knoxville (2025 and 2026) He conducted this research as an ORISE summer intern at ORNL (2023).

Deniz Aykac received a BS in Physics from the Bogazici University (1994) and an MS in Biomedical Engineering from The University of Iowa (2000). She has worked at Oak Ridge National Laboratory since 2002 extensively on 3D medical image processing, image, and video analysis.

Lauren Torkelson received a BS in Cyber Operations from Dakota State University (2019) and a MS in Cyber Defense from Dakota State University (2020). She has worked at Oak Ridge National Laboratory since 2021 and has worked on several research projects involving machine learning.

Thomas P Karnowski received a BSEE and PhD from the University of Tennessee (1988 and 2010), and a MS from NC State University (1990). He has worked at ORNL since 1990 on signal and image processing tasks.