– Independent Work Report Spring 2021 –

Privacy on the Go: Components and Challenges of License Plate Privacy in Video

Chandler Ault
Adviser: Dr. Olga Russakovsky

April 15, 2021

Abstract

As the number of surveillance cameras grow around the world, license plate privacy is becoming a cause for concern. This paper outlines the key components of a license plate privacy system and provides an implementation to act as a template for future work. Additionally, this paper outlines certain challenges in video privacy that became clear during the implementation. Namely, the challenges of dataset selection, overtraining, and selection bias.

1. Introduction

1.1. Motivation and Goal

Over the past few decades, the U.S. has seen an increasing number of cameras put in place for protection and surveillance. This trend saw the U.S. drop its population per camera ratio from 6.9 to 4.6 between 2015 to 2018; this ratio is nearly on par with China whose population to camera ratio was 4.1 in 2018 [16]. Although to most the possibility of a dystopian surveillance state today seems outrageous, there is cause for concern. This concern is even more alarming for minorities who are discriminated against every day without mass surveillance. When one considers how the growing power of modern computing could be leveraged against them, it is not hard to see how dire the circumstances are.

Although the possibility of a future surveillance state is real, that does not mean there is no abuse of the surveillance systems now. In fact, there have been a variety of abuses already reported

in America. The first form this takes is personal abuse. There have been numerous reports of police using cameras and license plate databases to stalk estranged spouses, to solicit strangers for romantic pursuits, to threaten motorists after traffic stops, or to do a variety of other inappropriate actions [22, 23]. Furthermore, and perhaps more difficult to measure, is institutional abuse. In the wake of the George Floyd protests, the then Attorney General Bill Barr increased the surveillance authorities of law enforcement to fight against drugs [7]. Not only is this likely to be abused on racial grounds since black people are much more likely to be arrested on drug related charges than white people, but it also posed the risk of silencing protesters by targeting them with the increased surveillance [7, 21, 24]. Additionally, there have been reports that traffic cameras and police cameras are more prevalent in black neighborhoods causing more fines and investigations being initiated in this minority group [25, 7]. Finally, there have been reports of unabashed criminal abuses. For example, one police official was caught gathering license plates of cars in gay club parking lots and blackmail married individuals[22]. Not only is this example criminal, but it is once again facilitating the discrimination of a minority group. Thus, there are many tangible concerns about the growing surveillance system in the United States.

All of these examples are in the United States alone, however, the possible future of a surveillance state is being realized in China today. Over the past few years there have been reports of China deploying a social credit score to incentivize "good behavior" in its citizens [10, 6, 14]. The social credit score is based on numerous factors, but some of these factors include jaywalking, buying or playing too many video games, buying Chinese made goods, and many more trivial factors. Although the consequences of a low social credit scores is sometimes as light as longer wait times at government institutions or jaywalkers being shamed by having their face displayed on public televisions, there are other cases in which a low social credit score can prevent someone from performing basic functions of life [14, 6, 10]. Some of the more dire repercussions of a low score include being barred from purchasing plane tickets, getting loans, renting a place to live, and more [14]. Seeing the power of a surveillance state in action in China should be eye opening as it

demonstrates how surveillance could be abused in other western countries such as the United States.

Briefly digressing into conjecture, as the COVID-19 pandemic rages on, it is not difficult to envision a world where a large-scale surveillance system could be used encroach on freedom in the name of safety. Consider a government tracking license plates and lowering one's social credit score or preventing them from entering essential stores for visiting a friend on their birthday or attending a large protest. This could easily be justified in the name of the public's health and safety, but it would clearly be a violation of the First Amendment. Furthermore, it would be prone to abuse as the government could attempt to silence the people on the most important topics such as racial equality.

Many of the most prevalent modern-day concerns of a surveillance state revolve around facial recognition software. Accordingly, most privacy addressing solutions to these concerns revolve around curtailing the use of biomarkers such as facial recognition. There are numerous face deconstruction and obfuscation frameworks in development some of which seek to put the decision of whether facial markers can be used in the hands of the individuals [2, 4]. Despite the work in facial privacy in response to the increase in cameras, there are noticeably less solutions addressing the problem of license plate detection and recognition in surveillance systems. Although it is evident that face recognition poses privacy concerns, it is equally true vis-à-vis license plate recognition since a plate is a one-to-one mapping to the owner of a vehicle. As previously detailed, this mapping of a license plate to an individual is frequently abused by law enforcement in various, and often discriminatory, ways. Due to this risk of abuse as well as lack of research into license plate privacy, my goal for this project was to create a template for what a real-time license plate privacy system may look like including license plate detection and obfuscation in a real-world setting.

1.2. Overview of Challenge

The key challenge of a license plate privacy system in the context of surveillance systems is to allow real-time, highly accurate privacy implementation. This is key because anything less than high accuracy will result in privacy being compromised. Additionally, the lack of real-time performance will reduce deployment opportunities since many surveillance systems have nearly endless security footage which would all require being processed. This would lead to backups and would likely not be practical to deploy the privacy system at all. Furthermore, often footage is being monitored by security guards. In an ideal world, this footage could be processed in real-time and provide license plate privacy before even reaching security guards on site.

Another important component of a successful license plate privacy system is the reversibility of the obfuscation. This is due to the fact that increases in surveillance cameras are often masked behind the cause of increased security and safety. Since there is evidence that criminals can be apprehended through license plate detection in security cameras, it is important to allow officials to retrieve license plates of vehicles provided this retrieval is done in a transparent and legal way such as with a warrant [5]. Without allowing retrieval of original footage, the system will not be practical for law enforcement or security applications, so the reversibility of the privacy system is an important challenge to overcome for practicality.

1.3. Overview of Previous Work

Since there is very little work in the area of license plate privacy, it is often beneficial to draw from work focusing on individual components of the privacy system. The first step is accurate, real-time license plate detection. Fortunately, the task of the detection and recognition of license plates is a robust field. There are various systems that tout quick runtime and accurate results, but the ones which marry these together the best are those implementing You Only Look Once (YOLO) models for detection.

Regarding providing privacy to license plates, the research is much scarcer. For this reason, one can look to facial privacy for systems for possible solutions. When reviewing the literature, two methods catch the eye. The first is to use the eigenvectors of a covariance matrix of a zero mean matrix comprised of a license plate image set to de-identify a targeted LP Preservative. More on this later. Another simpler method is to perform a pixel shuffle to obfuscate the license plate [4]. The benefit of this is that it is a simple, reversible implementation that renders the license plate unreadable to the human eye and character recognition systems.

1.4. Summary of Approach

My proposed approach is to implement all the necessary features. Namely, a detection model and an obfuscation method. However, since the goal is to have license plates obfuscated in every frame, an object tracking feature would cover frames when a license plate goes from detected to undetected suddenly due to a failure of the model. Additionally, to provide individualization and reversibility, a character recognition function should be utilized to tie a license plate to a single, unique key for reversal.

Considering the related work, my license plate privacy system utilizes a YOLO model for detection and a shuffle-based method for obfuscation. Additionally, due to seeking increased accuracy, I use Scale-invariant Feature Transform (SIFT) feature tracking to cover license plates that may go undetected in a given frame. These components form the base of my proposed approach. Character recognition, which is less critical in forming a template for license plate privacy, is solved by utilizing an existing package such as PyTesseract. However, since this paper's focus is on creating a template of a license plate privacy system, I will not address or consider the performance of the character recognition component of the system.

1.5. Summary of Results

Overall, the results of the system were good, but even a highly accurate system is often not sufficient in privacy since it requires near perfect results. With this being said, my model performed better than other comparable models in license plate detection on test video and did so in real-time. Furthermore, the addition of pixel shuffling and feature tracking rendered the license plate unreadable over 90% of the time.

The proposed system's license plate detection performs very well across all test data as well achieving precision, recall, and mAP.5 over 0.90 for all test sets. Additionally, the license plate tracking component of the system demonstrably improves the system's privacy by catching a significant portion of missed license plates. Finally, the obfuscation process proved to be qualitatively sufficient and reversible despite some minor artifacts due to resizing.

2. Problem Background and Related Work

The related work on this topic is multi-faceted since most related work focuses on only a small component of license plate privacy. For this reason, each of these components will be addressed individually in the following sections.

2.1. Automatic License Plate Detection (ALPD)

In the history of ALPD, there have been a variety of license plate detection methods suggested. Originally, these detection methods sometimes took the form of utilizing Histogram of Oriented Gradients (HOG) features descriptors, shape descriptors, color-based approaches and more [17, 19]. However, all these methods have certain drawbacks. Namely, these systems tend to perform poorly in conditions that the license plate is deformed, the lighting is poor, or the images are blurry [17]. For this reason, machine learning methods such as convolutional neural networks (CNN) are often employed. Although the use of a CNN meets the desire of accurate performance under a variety of

conditions, it simply cannot be deployed in real-time.

To address this, some real-time implementations of ALPD utilize YOLO models which boast incredible inference time and competitive mAP on the COCO dataset [13, 19, 9, 11]. The YOLO method has been further bolstered in recent years with the with the highly accurate YOLOv4 and the easily trainable YOLOv5 [15]. To see the difference in training time between these two versions, see Figure 1. These methods are allowing real-time ALPD at a rate in the range of 100+ FPS, and without a doubt are the future of real-time object detection.

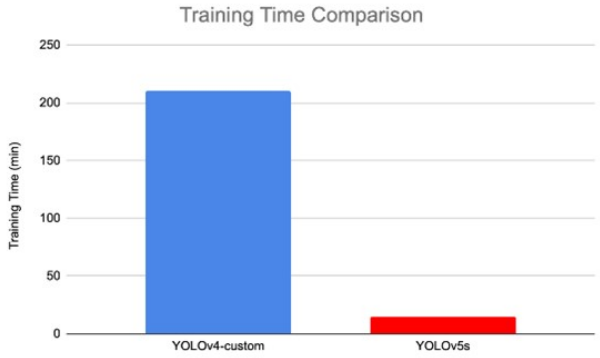


Figure 1: YOLOv4 and YOLOv5 Training Comparison

There are, however, two problems with the YOLO based methods. First, these methods focus on frontal views of license plates because they allow for easier character recognition [13]. Due to this, the training data is often not representative of difficult real-world conditions. On the other hand, one implementation sought to address this by tailoring a dataset of images captured by video from a car driving through traffic. However, simply utilizing real-world training data is not sufficient when considering license plate privacy. This brings us to the second problem which is that the majority of real-time YOLO-based ALPD methods do not consider implementation in actual video footage but opt to focus on the speed of analyzing a set of unrelated or loosely related images instead. My ALPD system would address these two problems by first training on datasets that represent real-life, poor-quality scenarios, and then employing feature tracking between frames of a video to focus on

providing highly accurate detection in video footage rather than solely in image datasets.

2.2. Privacy

Although there are many ALPD models, few of those seek to provide privacy in addition to detection. There are, however, some methods that address privacy implementation in license plates and facial features. One such privacy implementation for license plates is the use of inhomogeneous principal component blur (IPCB) [8]. In essence, this method takes the mean of a license plate image set and creates a zero-mean image set by subtracting the mean image from each image in the original set. Next, this method solves for the eigenvectors of the covariance matrix of the zero-mean image set. Finally, with these eigenvectors, projection coefficients, and probability distribution of what pixels represent sensitive data, they can create a de-identified license plate that appears to be untouched but has sensitive information with too much noise to read.

Another proposed method for de-identification is the process of secret block-based obfuscation [4]. In this method, the pixels of an image of a face are mapped into 8x8 square blocks, and a binary string with length equal to the number of blocks is randomly generated to act as a key. Each bit in the key will map to a block so that in the obfuscation stage, blocks corresponding to adjacent 1's and 0's can exchange key information. In the method proposed by Bo et al, only the AC coefficients are swapped for aesthetic reasons [4].

Although the IPCB method has very good qualitative results and is irrecoverable without possessing the coefficients of each principal component which is ideal from a privacy viewpoint, the image set in the implementation is required to be in grayscale to reduce its dimensionality which is not ideal for video privacy since ideally color would be preserved. Furthermore, the use of a license plate image set to find the eigenvectors requires detected license plates to be of a uniform type for IPCB to be effective. This reduces the generalizability of the system to various countries

with various types of license plates.

On the other hand, secret block-based obfuscation also has good qualitative results with slightly less secure privacy. This method is sufficient to prevent an individual or model from reading a license plate and provide total reversibility, but it is likely unable to withstand nefarious attacks due the limits of block swapping. Furthermore, this method was not applied to license plates which will make it novel in the realm of license plate privacy.

3. Approach

I approached the license plate privacy system different than other detection and obfuscation methods by placing the focus on its application in real-world videos. This focus required three key components: license plate detection, license plate tracking, and license plate obfuscation.

3.1. Detection

Foremost my system aims to be run in real-time, so my approach is to utilize a YOLOv5 model to achieve this. Most ALPD systems are built from the YOLOv3 models, so utilizing a YOLOv5 model should see increased performance in training time, runtime, and accuracy over older models (Figure 2) [20].

Model	AP^{val}	AP^{test}	AP_{50}	$Speed_{GPU}$	FPS_{GPU}		params	FLOPS
YOLOv5s	36.6	36.6	55.8	**2.1ms**	**476**		7.5M	13.2B
YOLOv5m	43.4	43.4	62.4	3.0ms	333		21.8M	39.4B
YOLOv5l	46.6	46.7	65.4	3.9ms	256		47.8M	88.1B
YOLOv5x	**48.4**	**48.4**	**66.9**	6.1ms	164		89.0M	166.4B
YOLOv3-SPP	45.6	45.5	65.2	4.5ms	222		63.0M	118.0B

Figure 2: Speed and Accuracy Comparison (YOLOv5 vs YOLOv3) [20]

Additionally, my system aims to perform well in license plate detection under difficult conditions by carefully selecting the datasets it is trained with. Since most datasets contain large, clear, centralized license plates, many models perform poorly under difficult conditions where the license plate is small, blurry, or partially covered. My approach to mitigate this poor performance was to initially train the model with a dataset containing clear plates to give high confidence in such settings, and then fine tune the model with a dataset containing more difficult images.

3.2. Tracking

Since the frames of a video should be related to each other in a continuous way, the ability to track a license plate instance across frames can be employed to increase privacy. This is perhaps the key novel idea in this license plate privacy system since most ALPD systems are not particularly concerned with going above and beyond the model to ensure that a license plate is detected in every frame. Additionally, most other systems do no seek to leverage the continuity of video frames to improve accuracy in their systems. My approach is to use SIFT features to do this since OpenCV has easy and fast packages to detect features in separate images. Furthermore, SIFT features employ many of the same non-machine learning techniques detailed earlier, so its use acts as a second layer of license plate detection beyond the YOLOv5 model.

3.3. Privacy

My method necessitates generalizability since it will be trained and tested on image sets containing various license plates. For this reason, IPCB will not work for my license plate privacy system. Secret block-based obfuscation will, however, be a very practical placeholder. Since this system is intended to be a template, block-based obfuscation is sufficient. That being said, in an actual deployment of a license plate privacy system, the methods for privacy can be very individualized to the country that the license plates will be coming from. Thus, IPCB would be worth attempting to

employ in such systems.

However, given certain limitations to the uniformity of datasets, secret block-based obfuscation is the method of choice for providing privacy for this system. Although the system in Bo et al. swap AC coefficients of adjacent blocks for aesthetic reasons, my system will not do this since by definition a license plate is blocklike to begin with. Another benefit is that this method will provide reversibility since a applying the same key in the shuffle of boxes will reverse the obfuscation. Finally, a single key can be applied to a single instance of a license plate across many frames which will give individualization to this privacy method.

4. Implementation

The objective of this system is to provide privacy in real-life video settings. Thus, the input to the system can be assumed to be either traffic or security footage. After this passes through the system, the resulting output should be the same footage with the privacy compromising license plates obfuscated. The following sections will detail how my system went about achieving this goal.

4.1. Architecture

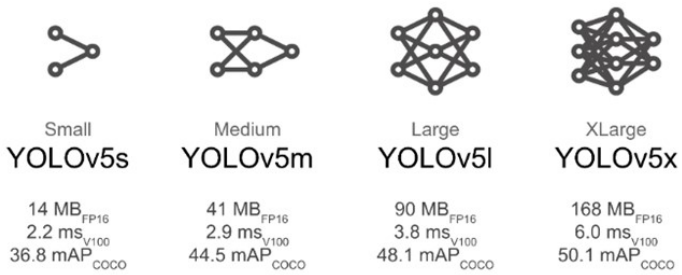


Figure 3: Comparison of Provided YOLOv5 Networks[1]

As previously stated, the detection model chosen for this system is YOLOv5. YOLOv5 comes with various sizes of base networks, and when choosing one, it is important to consider the attributes of the system and which model bests contributes to these. For a real-time privacy system that

implements license-plate tracking and obfuscation on every frame to only a single object-type, one of the smaller networks is appropriate. As the networks shrink in size, their inference time also shrinks. Additionally, the fact that there is only one object to be labeled per frame (i.e. license plates), larger networks would utilize an unnecessary amount of storage space. Thus, this system employs the YOLOv5s network which is the smallest model and claims to yield an inference time of 2.2 ms. per image [1]. The network also requires very little storage space. Overall, the YOLOv5s is about 10x smaller than the largest model size and almost 3 times faster (Figure 3).

4.2. Dataset

Selecting a dataset was one of the most difficult parts of the setup process. For the first round of training, the model was trained on a subsection of the Chinese City Parking Dataset (CCPD) [26]. This dataset contains over 250k car images with a single license plate visible. CCPD also primarily contains unobstructed license plates though it overcomes this by containing poor lighting and blurry conditions in some of their images (Figure 4). Due to space and computational limitations, my model was trained on a subsection with a size of about 5,000 images.

The fundamental flaw of the CCPD is that it does not represent difficult real-life scenarios incredibly well. To combat this, a second, smaller dataset was used to improve detection in these settings. This dataset is the Romanian Dataset of License Plates (RDLP) [12]. This dataset had a little over 500 training images with numerous plates per image many of which were partially obstructed or blurred. Additionally, this dataset possesses images under a variety of lighting conditions including night time photos.

4.3. Detection and Recognition

For detection, the YOLOv5s model was trained on both datasets. Initially, the model was trained on the CCPD for 100 epochs, and the weights that performed the best on the validation dataset

were saved. YOLOv5 models train very quickly, so only marginal improvements would be made by training for much longer, and as can be seen in the figure, there was a general upward trend in precision, but a general downward trend in recall as the number of epochs grew (Figure 5). For privacy, it is ideal to have increased recall since a higher value in this statistic indicates fewer false negatives which are catastrophic for privacy.

After training the model on CCPD, it was then finetuned on the RDLP to accommodate more

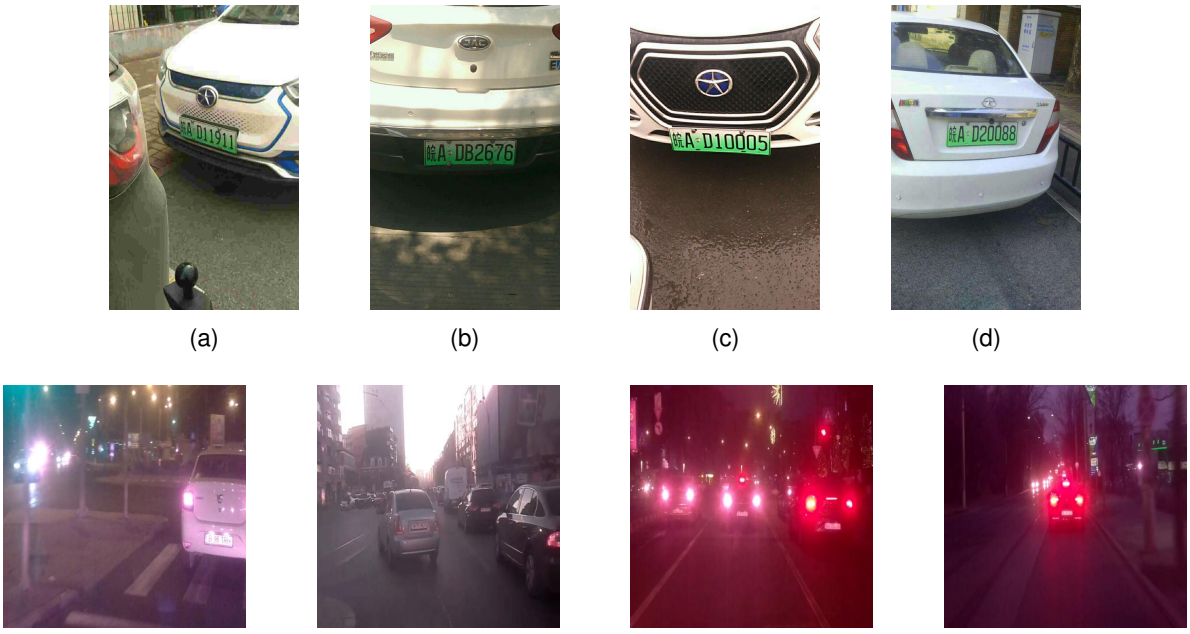


Figure 4: Examples of CCPD (a, b, c, d) and RDLP (e, f, g, h) [26, 12]

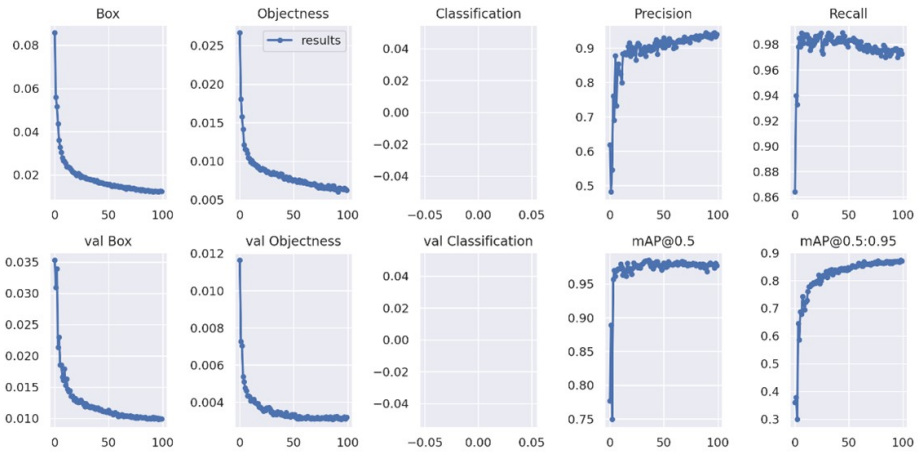


Figure 5: Training Graphs on CCPD

difficult scenarios. To do this, the weights from the CCPD stage of training were used as a starting point in the RDLP round of training. Since the RDLP is a much smaller dataset, data augmentation was utilized to not only increase the number of images, but also to give a more diverse set of images with obstructed license plates. Each image had three augmentations of a random cropping between 20-60% of the images' dimensions. By combining CCPD and RDLP, the objective is to improve the detection accuracy so that it is greater than a model solely trained on ideal plates. We can measure this effect by comparing the iteration of the model that was solely trained on CCPD with the iteration trained on both the CCPD and the RDLP datasets.

A secondary component in the system is the license plate recognition feature. This feature is not crucial for creating a template of license plate privacy, but it would be necessary for an actual instance of deployment. Recognizing the license plate allows the shuffle key to be stored along with the license plate number in an encrypted way that would require a private key associated with the license plate number to decrypt.

My system implements a license plate recognition protocol that leverages the fact that a license plate will be visible across many frames in a video. After detecting the region of the license plate, this section of the frame is passed to an optical character recognition (OCR) package called PyTesseract to read the plate. Although fairly accurate in many cases, PyTesseract is not sufficiently accurate frame-to-frame in a real-world video setting. To overcome this, each unique output from PyTesseract is saved and a tally is kept of the number of times each output is seen. After it has been predicted that a license plate has left video's field of view, the most common saved iteration of this instance is outputted as the license plate's predicted number. This method requires video to be successful in outputting the correct license plate numbers. Since annotated video data is difficult to come by and the video that will be used in testing later on only has a few unique license plates, there will be no measurement of this methods performance in license plate recognition.

4.4. License Plate Tracking

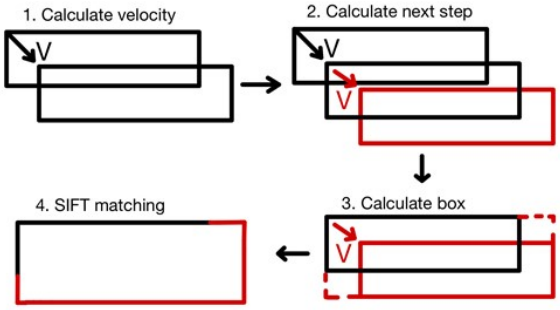


Figure 6: Tracking pipeline: (1) Calculate Velocity of the license plate across previous frames. (2) When license plate becomes undetected, calculate approximationg of its location using velocity. (3) Calculate a box to search for license plate features by taking the extreme coordinate values of the last known location and the approximated box. (4) Perform SIFT matching in this box.

The tracking component of my privacy system is the most unique aspect compared to other literature. This feature has a few key steps. First, when a new license plate is detected, the region of detection has its SIFT features extracted and stored in the license plate data structure. As this license plate remains detected and moves through the video, its velocity is calculated by taking the average number of pixels the center of the license plate traversed per frame in the previous 30 frames. If the detection model suddenly fails to detect the license plate in a frame, an estimated region in which the license plate could be is computed by adding the velocity coordinates to the last known location of the license plate. The maximum and minimum x and y coordinates of the corners of the last known location and the predicted location form the region in which a SIFT matching is attempted. This process can be seen in Figure 6.

If this matching returns a sufficient number of similar key points between the features stored in the license plate data structure and the new region, then the area encapsulating these key points is obfuscated as if that specific license plate was detected there. If there are not a sufficient number of key points, then the process repeats with a growing search region. After a certain number of frames in which there have been an insufficient number of key points, the instance of that license plate is

removed, and the shuffle key and plate number are logged. The efficacy of this implementation can be measured by the how many instances of missed license plates are caught by the addition of license plate tracking.

4.5. License Plate Privacy

The privacy component of the system is a fairly simple but effective approach. When a new license plate enters the video, a new, random binary shuffle key with length of 120 is generated and stored in the license plate data structure. Subsequently, for each license plate in each frame, the detected region of the license plate and the associated shuffle key are passed to the shuffling function. This function scales the license plate to a 48x160 pixel image. This rescaling allows the image to be partitioned into 120 8x8 blocks. This way, each block has a corresponding bit in the shuffle key. Now, the blocks associated with 1's are swapped with adjacent blocks associated with 1's, and the same is true of blocks associated with 0's (see Figure 7). The shuffle key should be the same across every frame of a license plate, and by simply recalling the shuffle function the obfuscated license plate with its associated key, the pixels will be unshuffled. This is a qualitative approach, so its evaluation will be of the same nature.

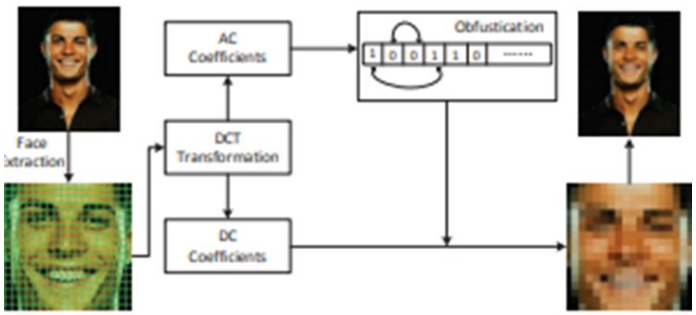


Figure 7: Secret Block-Based Obfuscation Summary[4]

5. Evaluation

5.1. Experiment Design

There are a couple components to consider when evaluating this system. The first and possibly most important component is the efficacy of the license plate detection itself. Two datasets were used in training, and both of these datasets had an accompanying test subset set aside for the purpose of evaluation. Additionally, to test the efficacy of the system's detection in video footage, four 30 second videos were manually annotated using VATIC. All four videos are from the same security camera footage. The videos then had each of their frames saved as an image, and each frame's annotation was subsequently converted to YOLO format. Thus, to evaluate the efficacy of detection, test sets from CCPD and RDLP will be used as well as the short video frames.

To gauge the performance of my model on these datasets relative to other implementations, I will include the performance of other YOLOv5 models as well. The first of these models (ANRP) was trained on an image set similar to that found in CCPD, but with far fewer images [3]. This model was created for use as a comparison to YOLOv3 models. Second, I will compare my results to that of another model (Adverse Environments) trained on CCPD as well as on other hand-annotated video footage [18]. This model claims to be trained for adverse environments, and since it was trained on CCPD as well, it should have similar performance to my model. The difference will lie in the selection of images with adverse conditions and how those conditions manifest.

The second component that will be evaluated is the system's license plate tracking feature. This will only be evaluated in the video data since license plates cannot be tracked across non-related images. To evaluate its efficacy, I will take $\frac{\#LP's\ Caught\ By\ Tracker}{\#LP's\ Missed\ By\ Detector}$. This will yield the rate at which the tracking system picks up on license plates the detection model misses. In essence, this will be the efficacy of the failsafe.

Finally, the license plate obfuscation will be solely measured in a qualitative manner on video footage. It is difficult to quantify the effectiveness of a privacy method, and it should be viewed as a qualitatively effective placeholder. Furthermore, I will demonstrate the quality of not only the obfuscation, but the reversibility as well.

5.2. Metrics

The metrics that will be considered when measuring this system's detection efficacy are precision, recall, and mAP.5. Precision and recall are both good metrics, but recall is certainly the most crucial in a privacy setting, so particular interest will be placed in raising recall. As previously stated, a higher recall means a higher rate of the relevant elements (i.e. license plates) were detected. On the other hand, precision indicates the rate at which selected items are relevant. If one is looking to maximize privacy, false positives are of far less consequence than false negatives, and it is for this reason that recall will be the key metric. Lastly, using mean average precision is useful to demonstrate the models overall efficacy.

5.3. Results and Comparison

Table 1: Performance Comparison

Dataset	**CCPD**	**RDLP**	**Video**
Model	P | R | mAP@.5	P | R | mAP@.5	P | R | mAP@.5
ANRP	0.496 | 0.172 | 0.143	0.596 | 0.172 | 0.143	0.063 | 0.067 | 0.041
Adverse Environment	0.910 | 0.401 | 0.630	0.784 | 0.909 | 0.919	0.824 | 0.712 | 0.884
Proposed	0.949 | 0.974 | 0.982	0.796 | 0.609 | 0.616	0.964 | 0.772 | 0.780
Proposed + Finetuning	0.919 | 0.925 | 0.949	0.935 | 0.906 | 0.954	0.975 | 0.955 | 0.967

Overall, the results were much of what was to be expected. ANRP performed poorly on nearly every test which is to be expected given the scope of its training data. This poor performance highlights the importance of a diverse and robust dataset for performance. More interesting, however, was the comparison between my proposed model and the Adverse Environment model. This model

held its own across all test datasets and even outperformed the proposed model in recall for the RDLP test. However, its performance was noticeably worse in the video test. Evidently, the Adverse Environment model had difficulty capturing larger license plates which is likely an artifact of its training data (Figure 8).

(a) ANPR (b) Adverse Environment

Figure 8: Example of the Adverse Environment testing results on larger and smaller license plates.

In light of the comparison of results, it is also useful to consider the comparison of each model's training data object distribution. As seen in Figure 9, each model had a wildly different distribution of the coordinates of its training data. In models trained with a wider distribution of coordinate values, that is in the Adverse Environment and Proposed system with finetuning, we see better recall performance in RDLP and higher mAP.5 in the video test. In the ANRP model, we see poor performance in nearly every test set. It is not a coincidence that this model was trained on a very limited array of images. Thus, as scenes become more difficult, we see a general increase in performance for models with a larger training distribution for license plate coordinates.

Another factor to consider when training is the dimension diversity of the license plates in the training set. Here we see an enlightening trend in the tested models. As previously mentioned,

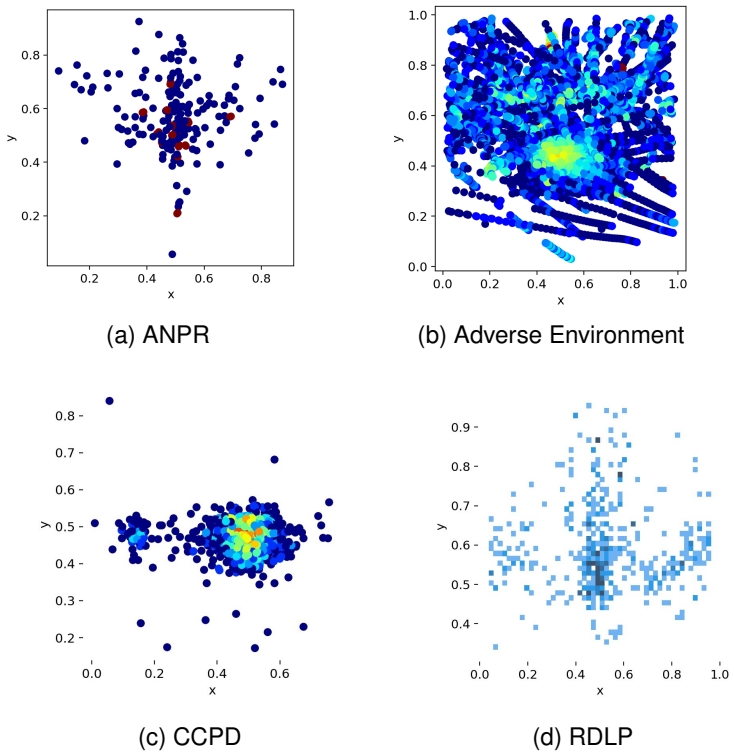
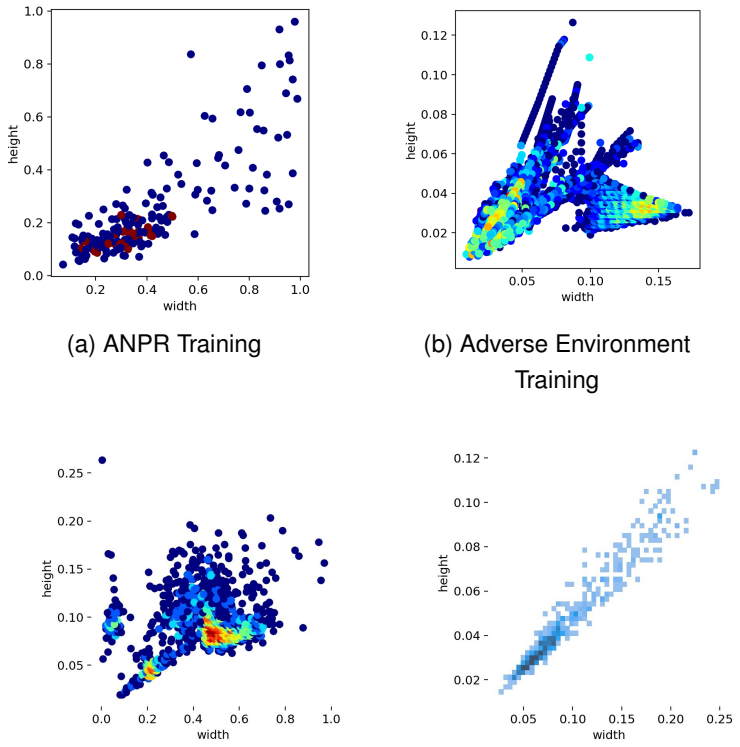


Figure 9: Distribution of license plates in the respective training sets. Measured in normalized image coordinates (Note different scales).

the Adverse Environment model qualitatively had more trouble on larger license plates. This is despite being trained on CCPD like the proposed model. One explanation is that there were 60% more self-annotated video images than CCPD images used in the model's training. One could conclude that this likely led to the model overtraining images in difficult conditions and specifically conditions with distant license plates. By observing the stats in Figure 10, we see that this could be the case. There is a strange distribution of the dimension data in the Adverse environment training data with nearly no plates exceeding 10% the dimensions of the training images. When compared to the other models, we see a similar trend in the RDLP data; however, this is likely compensated by the noticeably wider distribution of dimensions in the original model's training dataset.

This begs the question where the discrepancy comes from if both models utilized CCPD at least in part. One thing of note is that the Adverse Environment model contained about 60% more self-annotated video images than images from CCPD. However, given the distributions seen in



(a) ANPR Training

(b) Adverse Environment Training

(c) CCPD

(d) RDLP

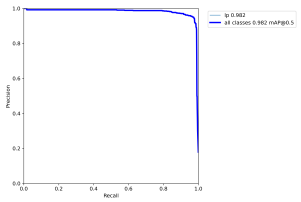
Figure 10: The height and width factors of license plates in the respective training sets. Measured in proportion of overall image height and width (Note different scales).

Figure 10, one would think that the Adverse Environment model would have contained at least some license plates exceeding the observed dimensions. One possible explanation is that in an attempt to train the data for difficult environments, a selection bias was introduced. Although the specific image set upon which the model was trained is not provided, an example of training labels (Figure 11) lends some credence to this theory.

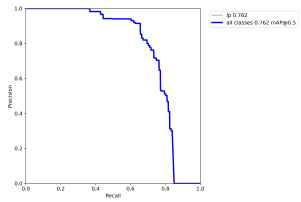
Overall, my model achieved a precision, recall, and mAP.5 all over 95% on the video test data. This is incredible performance and demonstrates that this model could nearly be deployed in a real-world surveillance setting. Perhaps one of the most interesting results was to see the difference in performance between the proposed model trained only on CCPD and the final model trained on both CCPD and RDLP. Despite a slight dip in performance of the final model on the CCPD test set, it achieved far better results as the conditions of the license plates became more unfavorable. The

Figure 11: Subsample of Adverse Environment predictions during training.

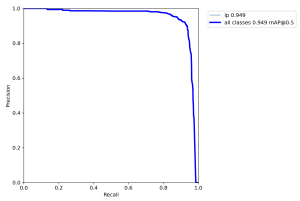
comparison of the original and finetuned models' performances on the CCPD and RDLP test sets can be seen in Figure 12.



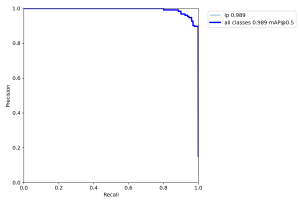
(a) Original, CCPD



(b) Original, RDLP

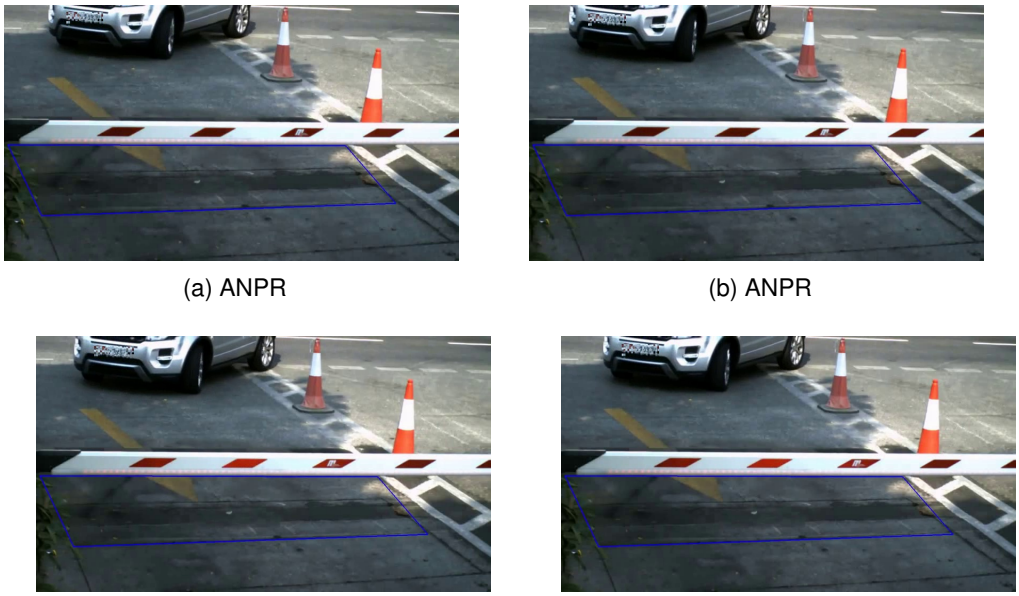


(c) Final, CCPD



(d) Final, RDLP

Figure 12: Pr curves for the original and final models on the two image test sets.



(a) ANPR (b) ANPR

(c) ANPR (d) ANPR

Figure 13: A sequence of frames successfully obfuscated using the tracking feature.

Now, it is time to measure the efficacy of the license plate tracking component of the system. Overall, the tracker performed adequately and caught 18 of 55 misses by the detection model in the video test set. This means that the tracker caught approximately 33% of the missed license plates (Figure 13). It is important to note here that the tracker will only be able to catch license plates that go from detected to undetected in the video. This is because it only searches for the SIFT features of a previously observed license plate. For this reason, the tracker would have slightly better results if just considering these examples rather than all missed detections. That being said, sometimes the tracker would blur the license plate without actually matching SIFT features accurately (Figure 14).

Overall, this is a very successful implementation of a tracker. Although it would ideally capture a much higher percentage of missed license plates, in the realm of privacy, every detection from tracking is important. This represents a fairly significant improvement in detection accuracy over the baseline model.

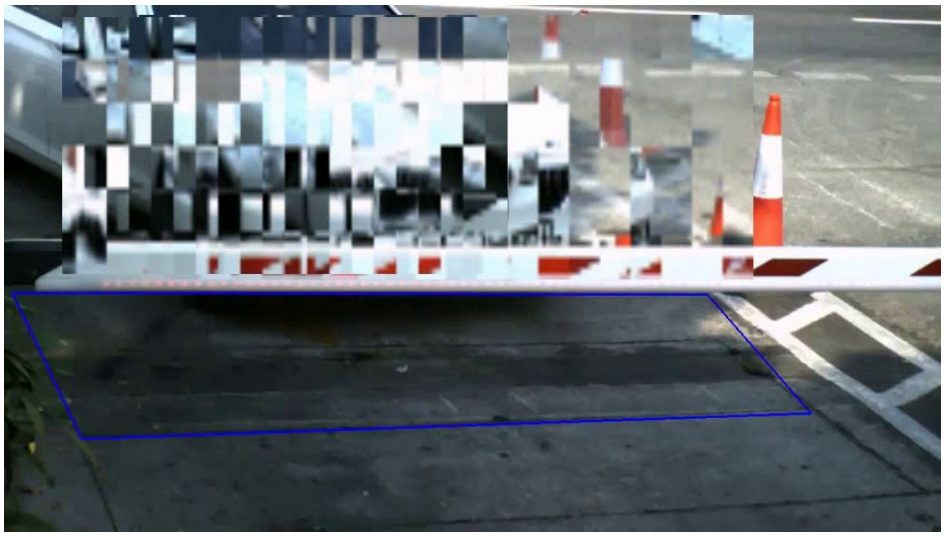


Figure 14: An example of a rare inaccurate SIFT matching.

5.4. Qualitative Results

Figure 15: Secret Block-Based Obfuscation on CCPD LP's

After reviewing the quantifiable results of the system and comparing them to other relevant models, it is time to consider the unquantifiable results of the secret block-based obfuscation component of the system. A few examples of this shuffling performed on CCPD images can be seen in Figure 15. Qualitatively, this shuffling performs very well as the license plates are unreadable for all intents and purposes. However, sometimes swapped boxes can be inferred by stark color changes as seen in some of the images. The obfuscation in the video test set was also of good quality, and they were arguably better than those in the CCPD due to the color scheme of the video test giving away less spatial information (Figure 16).

Another element of the privacy component of the system is its reversibility. Although the block-based shuffle method is easily reversible, some artifacts do arise when performing this function (Figure 17). As can be seen, the recovered license plate has a grid-like artifact introduced. This is due to the resizing performed on a detected license plate to get it into the dimensionality required to

(a) 0 Detected LP's, 1 Blurred (due to tracking)

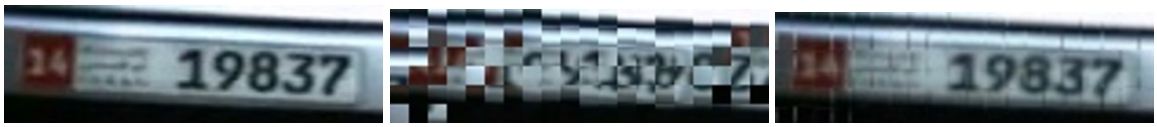
(b) 1 Detected, 1 Blurred

(c) 2 Detected, 2 Blurred

(d) 1 Detected, 1 Blurred

Figure 16: Proposed System's Blur in Video Detection.

divide it into 120 8x8 pixel boxes. Although this does not affect the readability of the recovered plate, it does present some issues in preserving video quality. Overall, however, the use of secret block-based obfuscation performs its roll well by rendering the license plate unreadable yet allowing it to still be recovered.



(a) Original (b) Obfuscated (c) Recovered

Figure 17: The three stages of license plate detection, obfuscation, and recovery of obfuscation.

6. Summary

6.1. Conclusion

Although far from perfect, my license plate privacy system highlights the most critical components for a viable privacy system, and it provides a template for how these components could fit together

in a real setting. These components include real-time, accurate detection in difficult scenes with a particular focus on detection in video; license plate tracking as a failsafe for the detection model; and a reversible and individualizable obfuscation method. With these three components, I believe additional, more robust systems could be developed and successfully deployed to address license plate privacy concerns. Additionally, in the scope of license plate detection itself, my implementation shows that care must be taken with what data is utilized to train an ALPD system as neglecting difficult conditions as well as focusing on them too much may lead to failure in real-life deployment. Finally, the use of object tracking in my system demonstrates that the fusion of a machine learning model and feature tracking can lead to improved performance.

It is important to stress that as it stands, this system would fail to provide privacy in many instances. In fact, I believe this failure illuminates the challenges of providing privacy in real-world video settings. First, the nature of privacy demands near perfect detection rates while the nature of video footage demands real-time inference rates. Often these two objectives are in direct conflict. Second, and related to the demands of near perfect detection rates, is the fact that an effective privacy system requires it to be highly individualized. This is true in tailoring a dataset as the most effective models will be solely trained on license plates from the specific country of its deployment. Additionally, this is true when considering what angles, obstructions, and lighting conditions the system may experience. Finally, this challenging requirement of individualization is true when considering what privacy method to utilize in the obfuscation of license plates. Although the IPCB method is cryptographically more secure than secret block-based obfuscation, it is not feasible to employ unless the training data and the testing data are both uniform. These challenges hold true in license plates, faces, and any other features one may want to provide privacy to. Thus, one key takeaway should be the challenges in providing privacy in video footage outlined here. By taking not of these challenges, others with aspirations of developing a privacy system will know what to expect.

Perhaps the most important conclusion to be drawn from this research project is in the analysis of

comparing this system's detection model to the others. While the models used for comparison are far from state of the art, they each highlight a component of training a model for video detection that should not be overlooked. First, it is necessary to have a sufficient amount of data in a variety of settings to avoid overtraining. This is an obvious observation, but it is one that can easily be seen in the ANPR model. Second, and perhaps more nuanced, is the necessity for a diverse dataset. This can be seen in the Adverse Environment model. Whether through selection bias or misfortune, the data used for training captured only distant, small plates. This led to poor external validity when applied to outside video datasets, and this indicates that the model will not be generalizable. When tailoring a dataset for real-world video settings, one should take care in selecting images to ensure diversity of dimensions and locations. Failing to do so may result in poor performance. Additionally, when using other datasets, one should understand the strengths and weaknesses of the dataset and compensate accordingly. For instance, the CCPD dataset had many clear, centered license plates, and to compensate for this, another dataset was used for finetuning. Thus, the key takeaway in this regard is to understand the datasets and to utilize caution when selecting training data.

My system was designed to highlight the necessary components of a privacy system and provide a template for future implementations. As cameras multiply and the United States and other countries march towards the possibility of experiencing abused surveillance power, it is important to be thinking of constructive ways to combat this. As previously detailed, license plate privacy is an important area which requires novel solutions. One day, I hope to see a privacy-affirming system that provides reversible, individualized privacy so that our neither our privacy nor security is compromised as long as said security is accompanied by the necessary legal documentation.

6.2. Limitations

There are a variety of limitations for my implementation. First, the sheer size of datasets was a limitation. Most datasets are annotated in VOX XML format which presents difficulties when training YOLO models because each annotation must be converted to the YOLOv5 annotation format before training or testing. This presents a limitation when using third party programs like Roboflow because the number of images that are allowed to be converted and stored are limited. Additionally, the lack of continuity between datasets presents a limitation especially if one seeks to implement the IPCB method. Finally, another limitation is the lack of literature attempting similar objectives. In time, likely more literature will appear, but as it stands there is nearly nothing addressing license plate privacy in real world video settings. This presents challenges when designing an outline of the system because there are no references.

6.3. Future Work

Since this system acts as a template for license plate privacy in video, there are numerous opportunities for future work. Namely, my system does not address encrypting and storing shuffle keys with their associated license plate number. This step would require the system to have highly accurate license plate recognition, to query a database of public keys associated with the plate number, and to use the public key to encrypt and store the shuffle key. This would also likely require some marker indicating which plate the key is associated with. Implementing this feature would round out the system and make it viable for security purposes. As it stands, the system only detects, recognizes, tracks, and blurs the license plate. The retrieval of already obfuscated license plates, though possible, is not implemented.

Additionally, and more compelling in my opinion, would be to focus on how data selection impacts performance in ALPD. From the minimal evidence laid out in this paper, there seems to be a correlation between performance and data diversity, but what are the quantifiable effects of

selecting only small license plates or only large, clear license plates? Furthermore, what would the ideal combination of difficult and easy conditions be in a license plate dataset? These are the questions that will likely need to be answered before a viable license plate privacy system will be able to be implemented since these are the questions that will determine the efficacy of a ALPD model. It is very possible that the answer to these questions will be generalizable to other object detection models as well which makes them excellent future things to consider and work on.

6.4. Note to Future Computer Vision IW Students

To any future students doing independent work in computer vision, I advise you to explore much and focus on little. In this IW class, you are asked to choose a topic very early on in the class before you have been fully exposed to many topics of computer vision. This will be one of the most important choices made for the next semester or even the entire school year. Thus, when choosing a topic, make sure to read and learn about as many computer vision topics as possible even on the most superficial level. By doing this, you will undoubtedly find a topic that excites you. While my project was enlightening, I believe there could have been other areas that would have been more engaging if I had known about them sooner. Do not let this happen to you.

Though I advise to learn a superficial level about much, I also advise you to focus on very little. One regret that I have in my independent work is that I feel like I had too grandiose aspirations for the project. Since this is likely one of your first opportunities to have this level of freedom in a project, it is better to narrow your scope and progressively widen it as you achieve your goals. As a single undergraduate student, it is unlikely that in a semester or two you will achieve grandiose aspirations. This is a realization I had to come to. However, often in those grandiose aspirations lies a myriad of interesting, full fledged IW topics. For instance, one of the most interesting components of writing this paper was in analyzing how data selection may have impacted the models. This could easily be a great topic for independent work, but it is squashed by the weight my project's scope.

Thus, to summarize my advice, explore as many topics as possible before choosing your topic, and understand your limitations when you do make your choice. By doing this, you will always be engaged but also make a small, more impactful contribution to the field.

References

[1] [Online]. Available: https://pytorch.org/hub/ultralytics_yolov5/

[2] P. Aditya *et al.*, "I-pic: A platform for privacy-compliant image capture," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 235–248. Available: https://doi.org/10.1145/2906388.2906412

[3] S. Baldota, "How yolov5 solved an ambiguity encountered by yolov3," Aug 2020. Available: https://towardsdatascience.com/indian-car-license-plate-detection-using-yolo-v5-ae2574578175

[4] C. Bo *et al.*, "Privacy.tag: Privacy concern expressed and respected," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 163–176. Available: https://doi.org/10.1145/2668332.2668339

[5] By, "The advantages of a license plate recognition (lpr) system," Apr 2018. Available: https://ngscinc.com/advantages-of-license-plate-recognition-systems/

[6] L. Casiano, "How china uses its massive surveillance apparatus to track its citizens, keep them in line," May 2020. Available: https://www.foxnews.com/world/china-massive-surveillance-apparatus-track-citizens

[7] K. Crockford, "Aclu news and commentary," Jun 2020. Available: https://www.aclu.org/news/privacy-technology/how-is-face-recognition-surveillance-technology-racist/

[8] L. Du and H. Ling, "Preservative license plate de-identification for privacy protection," in *2011 International Conference on Document Analysis and Recognition*, 2011, pp. 468–472.

[9] S. Khazaee *et al.*, "A real-time license plate detection method using a deep learning approach," in *Pattern Recognition and Artificial Intelligence*, Y. Lu *et al.*, Eds. Cham: Springer International Publishing, 2020, pp. 425–438.

[10] P. Knox, "China building spy state with emotion monitor cameras and 'social credit' scores," Mar 2021. Available: https://www.thesun.co.uk/news/14276151/china-plans-high-tech-streets-spy-cams-inform-neighbours/

[11] R. Laroca *et al.*, "A robust real-time automatic license plate recognition based on the yolo detector," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–10.

[12] R. Lucian, "License-plate-dataset," https://github.com/RobertLucian/license-plate-dataset, 2020.

[13] S. Montazzolli and C. Jung, "Real-time brazilian license plate detection and recognition using deep convolutional neural networks," in *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2017, pp. 55–62.

[14] M. Myers, "Here's the tech china's using to monitor, shame and rate citizens," Apr 2018. Available: https://www.cnet.com/news/china-turns-to-tech-to-monitor-shame-and-rate-citizens/

[15] J. Nelson, "Responding to the controversy about yolov5," Mar 2021. Available: https://blog.roboflow.com/yolov4-versus-yolov5/

[16] T. Ricker, "The us, like china, has about one surveillance camera for every four people, says report," Dec 2019. Available: https://www.theverge.com/2019/12/9/21002515/surveillance-cameras-globally-us-china-amount-citizens

[17] Z. Selmi, M. Ben Halima, and A. M. Alimi, "Deep learning system for automatic license plate detection and recognition," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 01, 2017, pp. 1132–1138.

[18] R. Shah, "Automatic-license-plate-detector-for-adverse-environments," https://github.com/THINK989/Automatic_License_Plate_Detector_for_Adverse_Environments, 2020.

[19] S. M. Silva and C. R. Jung, "License plate detection and recognition in unconstrained scenarios," in *2018 European Conference on Computer Vision (ECCV)*, Sep 2018, pp. 580–596.

[20] J. Solawetz, "Yolov5 new version - improvements and evaluation," Mar 2021. Available: https://blog.roboflow.com/yolov5-improvements-and-evaluation/

[21] unknown, "Race and the drug war." Available: https://drugpolicy.org/issues/race-and-drug-war

[22] unknown, "What's wrong with public video surveillance?" Mar 2002. Available: https://www.aclu.org/other/whats-wrong-public-video-surveillance

[23] unknown, "Police sometimes misuse confidential work databases for personal gain: Ap," Sep 2016. Available: https://www.cbsnews.com/news/police-sometimes-misuse-confidential-work-databases-for-personal-gain-ap/

[24] unknown, "Racial disparity in marijuana arrests," Mar 2021. Available: https://norml.org/marijuana/fact-sheets/racial-disparity-in-marijuana-arrests/

[25] T. H. . W. I. C. Writer, "Traffic cameras more prevalent in black d.c. neighborhoods," Jul 2018. Available: https://www.washingtoninformer.com/traffic-cameras-more-prevalent-in-black-d-c-neighborhoods/

[26] Z. Xu *et al.*, "Towards end-to-end license plate detection and recognition: A large dataset and baseline," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 255–271.