Association for Information Systems
AIS Electronic Library (AISeL)

Proceedings of the 2018 Pre-ICIS SIGDSA Symposium

Special Interest Group on Decision Support and Analytics (SIGDSA)

12-12-2018

Convolutional Neural Networks to Mitigate Transit Crowd Impacts

Jay Claiborne
Auburn University, jzc0019@tigermail.auburn.edu

J Locke
Auburn University, jml0062@tigermail.auburn.edu

Convolutional Neural Networks to Mitigate Transit Crowd Impacts

Research-in-Progress

J. Claiborne
Auburn University
jzc0019@tigermail.auburn.edu

J. Locke
Auburn University
jml0062@tigermail.auburn.edu

Abstract

Crowd counting has experienced an exponential evolution from mechanical turnstile counters to state-of-the-art computer vision applications. Crowd research stakeholders include designers, venue managers, public safety officers, school officials, religious organizers and more. An area of focus especially subject to crowding is public transit. Rail transit environments offer a number of unique hazards for pedestrians. This research seeks to mitigate the dangers of crowding in rail transit via the use of convolutional neural networks. By training a convolutional neural network (CNN) to count the riders on rail cars using existing CCTV cameras, this research hopes to inform commuters, so they can better position themselves on boarding platforms to avoid crowded cars. This research includes the development of a novel, public transit specific, dataset of internal rail car passenger images.

Keywords

CNN, Convolutional Neural Network, Crowds, Rail Cars, Transit

Introduction

Crowds are large groups occupying a single space and sharing a common goal (Forsyth, 1999). Pedestrian crowd dynamics have been studied from a variety of perspectives. Researchers have applied concepts from fluid dynamics (Hughes, 2003), gas kinetics (Pelechano & Malkawi, 2008), and particle movement theories (Mehran, Oyama, & Shah, 2009) in attempts to better understand and model the movement behaviors of pedestrians. Psychologists, sociologists, and physiologists have explored the complexities of both individual and group mental, social, and physiological responses to stimuli in crowd settings. A strong contributing factor to all of these studies has been the environment and its various influences and hazards (Bitner, 1992; Treuille, Cooper, & Popović, 2006). A particularly complex environment for pedestrian crowds is the public transit setting. Public transit facilities and passenger rail cars represent a unique convergence point of public safety and crowd dynamics concerns.

Attempts to alleviate crowding concerns have been addressed in a variety of ways. Automatic people counters have progressed from turnstiles and infrared light beam systems to complicated image analysis techniques. As a longtime staple in the safety/security discipline, the simple security camera can now add new value as computer vision drastically extends the utility of captured images. Computer vision presents an opportunity to elevate basic recorded video footage into higher value data useful for facial/object recognition, behavior analysis, crowd management, public space design, and people/passenger counting (Junior, Musse, & Jung, 2010). The primary goal of this study is to determine if the use of convolutional neural networks can be leveraged to mitigate the negative aspects of crowding in transit rail settings.

Crowd Impacts

Regardless of the environment, crowds are inherently unpredictable and dangerous. The physical dangers of crowds have been well documented. Between 1971 – 2011 there were a total of 156 mass incidents which included 3,332 fatalities and 4,508 injuries (Soomaroo & Murray, 2012). Some of the worst crowd disasters include:

- Hajj Mecca, Saudi Arabia (1990) – 1,426 deaths
- Johannesburg, South Africa (2001) -47 deaths
- Akashi, Japan (2001) – 160 deaths
- Hajj Mecca, Saudi Arabia (2006) – 346 deaths
- Duisburg, Germany (2010) – 21 deaths
- Phnom Penh, Cambodia (2010) -347 deaths

To appreciate the internal forces generated by surging crowds consider the following:

> *"It is difficult to describe the psychological and physiological pressures within crowds at maximum density. When crowd density equals the plan area of the human body, individual control is lost, as one becomes an involuntary part of the mass... Shock waves can be propagated through the mass sufficient to lift people off of their feet and propel them distances of 3 m (10 ft) or more. People may be literally lifted out of their shoes, and have clothing torn off... Crowd forces can reach levels that are almost impossible to resist or control. Virtually all crowd deaths are due to compressive asphyxia and not the "trampling" reported by the news media. Evidence of bent steel railings after several fatal crowd incidents show that forces of more than 4500 N (1,000 lbs.) occurred. Forces are due to pushing, and the domino effect of people leaning against each other."* (Fruin, 1993)

The risks, delays, and discomforts of dealing with overcrowding are well known and combine to modify pedestrian choices, behaviors, and perceptions. Negative customer experiences due to crowding include violations of personal space, violations of expected public behavior, feelings of loss of control/mobility etc. The physical safety concerns and social discomfitures combine to erode the perceived value of the transit service.

A study conducted on an Australian rail transit system found that factors representing available seating, crowding levels (moderate/heavy), and total time on the train were all significant and negatively correlated parameters for estimating perceived value (Collins & Hensher, 2017). In a Korean transit system, researchers found that 76.6% of survey participants intentionally chose a specific car. Of this group, 13.5% reporting selecting the car based on a pursuit of comfort (Kim, Kwon, Wu & Sohn, 2014).

Environment

Transit rail facilities present a number of physical safety challenges. Transit officials often find themselves faced with difficult choices between providing adequate space and managing costs. By necessity, facilities are often located in dense urban areas where premium real estate prices apply economic pressure to minimize the size of buildings. This financial pressure often finds itself at odds with the need to accommodate growing foot traffic. Large quantities of passengers are simultaneously moving into and out of facilities as trains constantly deliver and remove riders in large numbers. Inside, ticketing mechanisms create bottlenecks as time-conscious transit riders attempt to access boarding platforms. Fear of missing trains often results in customers overloading elevators, running on stairs or escalators, and pushing or jostling fellow riders.

Once customers reach the platform, the nature of time-interval departures can create significant crowds from seemingly moderate flows of pedestrians as they form queues between train departures. During large-scale public events, crowds at some stations can exceed the limits of safe passenger capacity on platforms. This is often true after sporting events, music concerts, and holiday festivities (New Year, 4th of July, etc.). Most rail transit systems operate without physical borders between passengers and potentially dangerous operational areas. Passengers are expected to remain within designated portions of the platform to avoid contact with moving trains or energized rail lines. If platforms become overly crowded (e.g. Figure 1) and incoming trains attempt to deposit more patrons into a station, customers can be inadvertently forced off double-sided platforms into the path of oncoming trains or into contact with high voltage rail lines.

During the 10-year span from 2007-2016, U.S. public transit heavy rail operations were responsible for 328 fatalities and 1,020 injuries. Injuries were defined as requiring immediate medical transportation away from the scene (Bureau of Labor Statistics, 2018). It is important to note that these totals only include reported injuries that required emergency medical staff to respond. It is reasonable to assume that many more injuries occurred but were not reported or failed to meet the Bureau of Labor Statistic's medical transportation criteria.

Rail cars offer similar if sometimes more concentrated forms of safety issues. Most rail cars are designed with 6 pairs of double doors for the loading and unloading of people, luggage, bicycles, wheelchairs etc. Commuters represent a wide range of ages, mobility capabilities, walking speeds, and familiarity levels with public transit. Because the rail car doors are typically situated opposite one another, double-sided platforms feed people from opposite directions into a four-way intersection of center aisles and entry/exit paths. This entry bottleneck often results in confusion and frustration as passengers attempt to find seats, stow large items, make their way into designated handicap areas, or claim standing space.

Modern passenger rail cars are equipped with a number of safety features. Two safety features dealing with the rail car doors are particularly relevant. Applying perpendicular force to train doors (pressure from the inside) triggers an alert that indicates to the operator that the doors are not in a safe position and that the train should not proceed. As crowding on trains grow to the extent that passengers are unable to refrain from being pressed against doors, the safety feature is activated and trains experience delays.

Rail car doors are typically open for approximately 8-10 seconds at each station. Densely packed trains may increase rider anxiety regarding their ability to exit in time (Chen, 2010). Crowded trains often prevent passengers from entering or exiting the train quickly enough to avoid being caught by closing doors. Typically, if the doors attempt to close and encounter an obstruction, the operator is notified not to proceed, and the doors cycle through a new open/close sequence. This interruption creates a minor system delay. A Washington D.C. transit authority reported that annual impact of delays from door closing issues amounted to 455 hours in single year (Hedgpeth, 2012).

In rare instances, the doors close on a customer, bag, or article of clothing without an accompanying notification to stop the train and restart the door cycle. This creates an especially dangerous situation wherein a commuter or a commuter's property is potentially dragged across the station platform. The potential for serious injury/damage to both the trapped person/object as well as other commuters on the platform increases significantly. A five-agency study of platform safety found that "Closed in Door" occurrences ranged from 4-25% of the total reported platform-train incidents (Hunter-Zaworski, 2017).

Automatic Passenger / People Counters

Retailers use people counters to better understand foot traffic, store browsing patterns, and marketing effectiveness (Perdikaki, Kesavan, & Swaminathan, 2012). Public safety researchers have adopted people counters as a way to supply crucial crowd pattern and behavior data (Junior, Musse, & Jung, 2010). By counting people and modeling travel patterns and flow rates, planners can design safer and more pedestrian friendly buildings and rail cars.

Automatic people counters (APCs) used in transit research focus on entry and exit counts for rail cars, gates, stations etc. Techniques for automatic counting include pressure mats/plates, infrared beams, thermal imaging, Wi-Fi signal counting, 3D stereoscopic sensors, laser counters, and computer vision systems. Each of these methods attempts to balance cost, maintainability, and capability to varying degrees. Simple systems like pressure mats and turnstiles are inexpensive to install and easy to maintain, but offer limited capabilities. More technical systems may require extensive installation and maintenance investments, but offer greater accuracy and extended capabilities.

Because they are external to the rail car, current counting methods have a number of shortcomings. They rely on single view systems that offer less than optimal accuracy rates. At most each rail car door will have one mounted counting apparatus. Platform mounted devices may not account for varying train lengths or situations where trains are not stopped in uniform locations relative to the platform. Lastly, external systems fail to account for passengers moving between trains after boarding. If commuters move to different rail cars during transit the density levels of individual cars could change significantly. In the case of an emergency or disturbance large numbers of riders might elect or be instructed to move between cars. Not knowing the exact location of riders are on the train could impede emergency medical or law enforcement efforts.

Computer Vision and Convolutional Neural Networks

Computer vision techniques offer advantages over traditional counting systems. In addition to accurate people counting, computer vision can be used for dangerous object recognition, and crowd behavior notification. CNNs have been used to achieve real-time gun recognition at a 93% accuracy rate (Verma & Dhillon, 2017), and to positively identify abnormalities in crowd behavior over time (Sabokrou, Fayyaz, Fathy, Moayed, & Klette, 2018). Many researchers have concluded that convolution neural networks have become the leading approach in computer vision research (Trottier, Gigu, & Chaib-draa, 2017). CNNs are easier to train while sacrificing little accuracy compared to more complex feedforward neural networks (Krizhevsky, Sutskever, & Hinton, 2012). This study proposes that the use of computer vision techniques, specifically convolutional neural networks, using the existing CCTV camera infrastructure, presents the most attractive option for transit authorities.

Research Questions

To earn consideration as the preferred choice of automatic people counters in rail cars, CNNs should be able to demonstrate equivalent or improved accuracy levels as compared to other methods currently in use. An affirmative finding regarding accuracy, coupled with the additional benefits of computer vision and use of existing CCTV camera infrastructures, would combine to make CNNs the APC method of choice.

1. Can convolutional neural networks be used to determine passenger rail car counts at an accuracy rate of at least 95%?

To test the effectiveness of using CNNs and the current CCTV camera network, it is proposed that stations display rail car passenger counts for inbound trains. This information will provide waiting commuters the opportunity to disperse along the platform in a matter that generally normalizes the distribution of passengers and/or takes advantage of available space on inbound rail cars. An existing relationship with a major U.S. transit authority presents the opportunity to monitor/record the platform distribution behavior.

2. Does displaying inbound rail car passenger counts cause occupancy rates of rail cars to normalize?

Related Work

In a review of deep learning, LeCun, Bengio, & Hinton (2015), described how CNNs were waning in popularity until the 2012 ImageNet competition. In a contest to classify approximately 1 million images into 1,000 different classes, CNNs were able to cut classification errors dramatically. The convergence of hardware (GPUs) and applications of techniques (RELU and dropout) reignited interest in CNNs, which are now considered the leading approach for most recognition and detection problems. They predicted that convolutional neural networks will be the impetus behind much of the future progress in computer vision (LeCun, Bengio, & Hinton, 2015).

Some researchers have gone to great lengths to develop new datasets to test their crowd counting models. A team of researchers from Shanghai Tech University collected and labeled almost 1,200 images with annotations for over 300,000 heads. They used this dataset to test a highly effective multi-column convolutional neural network (MCNN) to make crowd size predictions from a single image (Zhang, Zhou, Chen, Gao, & Ma, 2016).

Using a cross-scene technique, which does not require new annotations when new scenes or layouts are introduced, Zhang et. al focused on the dual learning objectives of crowd counting and crowd density (Zhang, Li, Wang, & Yang, 2015).

Focused on highly dense crowds, Boominathan et al. used a combination of CNNs to predict density maps. They found that multi-scale data augmentation, to account for wider variations in scale, helped inform their models (Boominathan, Kruthiventi, & Babu, 2016).

In a survey of crowd counting research, Lot et al. categorize crowd counting into three distinct groups: counting by detection, counting by clustering, and counting by regression. They also introduced a new dataset based on 60,000 shoppers pulled from 2,000 video frames of mall traffic (Loy, Chen, Gong, & Xiang, 2013).

A major consideration when training a neural network is the activation function which acts as an output from one layer and an input to the next. The most common activation function in use is the Rectified Linear Unit (ReLU). This is a half-rectified non-linear function (LeCun, Bengio, & Hinton, 2015). A related function, but with a wider range of possible values, is the Leaky ReLU (see Figure 4). Despite the common use of ReLU as the default activation function, in a comparison of activation functions for CNNs, variations of Leaky ReLU were found to produce smaller error rates and log loss values (Xu, Wang, Chen, & Li, 2015).

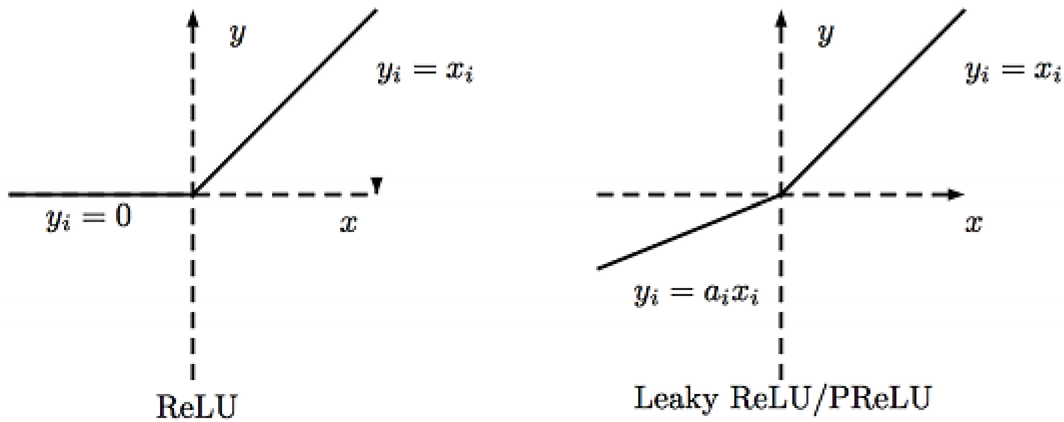

Figure 4. ReLu and Leaky ReLU

(Xu, Wang, Chen, & Li, 2015).

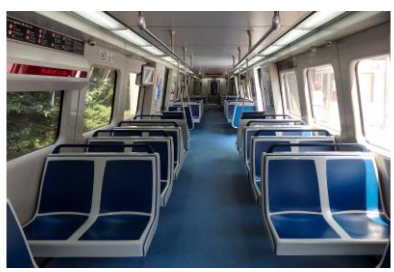

Figure 1. Crowded Platform **Figure 2. Rail Car View** **Figure 3. Rail Car Passengers**

Method / Model

This research will utilize a convolutional neural network to generate the passenger count of each rail car. Testing will include accuracy measures of single images (frontmost CCTV camera with lengthwise view, see Figures 2 & 3) and dual images (frontmost and rearmost CCTV cameras averaged). The selected activation functions for testing will be ReLU and Leaky ReLU.

Data

The data for this research will be sampled from daily operating footage of passenger rail cars. The participating system is a top 10 U.S. transportation authority and is highly representative of other transit systems in terms of passenger motivations, rail car layout, rider accommodations, and pedestrian safety concerns. Each rail car in the subject system is equipped with Wi-Fi and 10 ceiling mounted CCTV cameras (see Figure 5). Due to fixed camera positions and uniform rail car layouts, the effectiveness of convolutional neural networks is expected to be high. The initial dataset will consist of 500 still image pairs (forward most and rearmost views) captured from operational footage. The images will be reviewed, and ridership counts will be annotated for the training phase of the model.

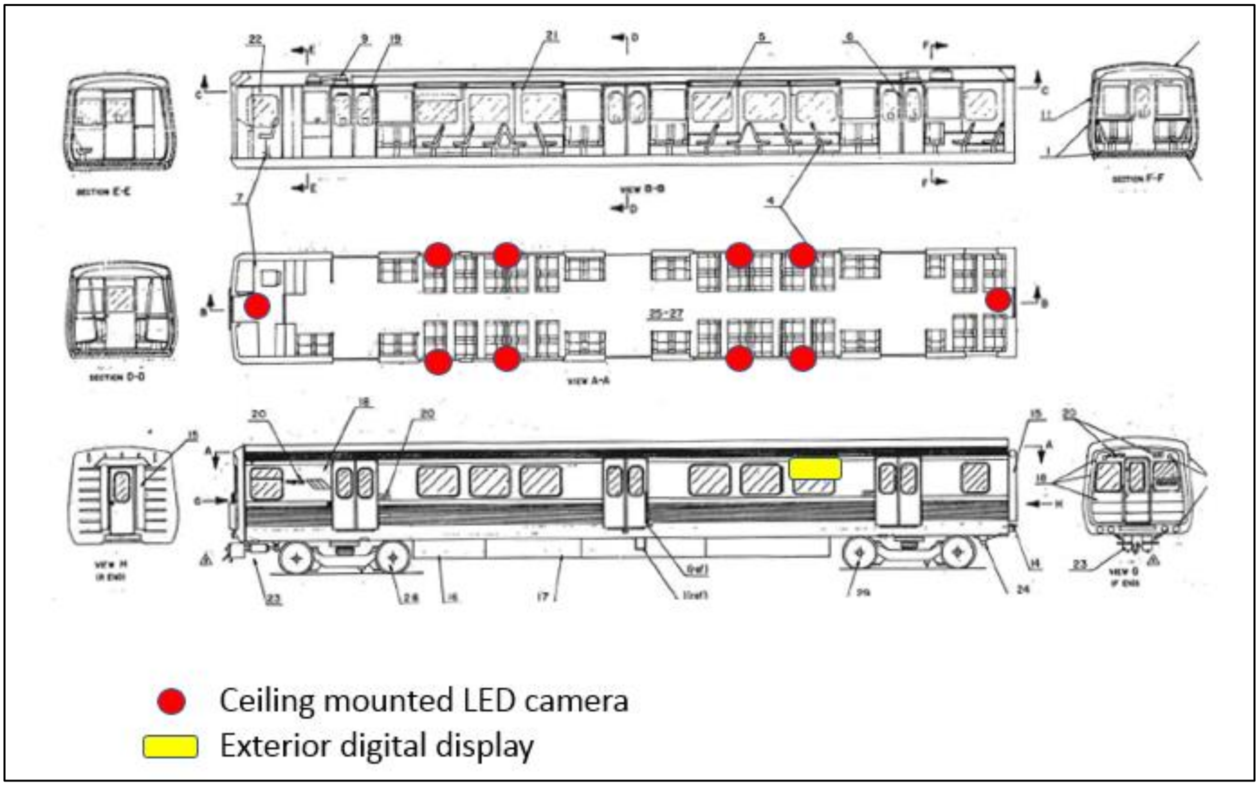

Figure 5. Rail Car Diagram

Limitations / Future Research

This research will initially only include one U.S. transit system. A 2017 study of preferred personal space revealed wide ranges in the tolerance for strangers in close proximity across nationalities. While Americans preferred to maintain approximately 95 centimeters between themselves and strangers, countries including Romania, Turkey, and Hungary preferred as much as 25% more space (Sorokowska et al. 2017). Cultural factors that influence pedestrian behavior could impact the effectiveness of displaying crowd density information.

Another limitation in this research is the dependency on either onboard hardware for local image processing, or network speed and reliability for remote image processing. If processing resources are limited, it may require design concessions that impact accuracy (channel reduction, image quality concessions, etc.).

Future research might include an investigation of the growth rate of cities with transit systems and the levels of tolerance for various negative features of crowding (e.g. violation of personal space or declines in perceived value of transit associated with crowding).

Conclusion

The benefits from this research are twofold. First, this study represents an opportunity to potentially mitigate the dangers of overcrowding in a transit setting. The dangers of crowding are well documented, and merit continued focus. Second, during the course of the research a novel dataset will be gathered and annotated depicting rail car passengers from multiple angles. This dataset can be used to assist in training and testing new transit related APCs.

References

Bitner, M. J. (1992). Servicescapes: The impact of physical surroundings on customers and employees. the Journal of Marketing, 57-71.

Boominathan, L., Kruthiventi, S. S., & Babu, R. V. (2016, October). Crowdnet: A deep convolutional network for dense crowd counting. In Proceedings of the 2016 ACM on Multimedia Conference (pp. 640-644). ACM.

Cheng, Y.-H., 2010. Exploring passenger anxiety associated with train travel. Transportation 37 (6), 875–896

Collins, A. T., & Hensher, D. A. (2017). Choice modelling with time-varying attributes, with an application to train crowding.

Forsyth, D. R. (1999). *Group Dynamics* (3 ed.). Belmont, CA: Wadsworth Publishing.

Fruin, J. J. (1993). The causes and prevention of crowd disasters. Engineering for crowd safety, 1(10), 99-108.

Hedgpeth, D. (2012, March 15). Frustrations over Metro train doors add up - for riders and operators. Retrieved September 15, 2018, from https://www.washingtonpost.com/local/trafficandcommuting/frustrations-over-metro-train-doors-add-up/2012/02/27/gIQAb5GQES_story.html?utm_term=.8cdb1d225d80

Hughes, R. L. (2003). The flow of human crowds. Annual review of fluid mechanics, 35(1), 169-182.

Hunter-Zaworski, K. (2017). Analysis of passenger incident data from five rail transit systems. *Safety*, 3(3), 21.

Junior, J. C. S. J., Musse, S. R., & Jung, C. R. (2010). Crowd analysis using computer vision techniques. *IEEE Signal Processing Magazine*, 27(5), 66-77.

Kim, H., Kwon, S., Wu, S. K., & Sohn, K. (2014). Why do passengers choose a specific car of a metro train during the morning peak hours?. Transportation research part A: policy and practice, 61, 249-258.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436.

Loy, C. C., Chen, K., Gong, S., & Xiang, T. (2013). Crowd counting and profiling: Methodology and evaluation. In Modeling, Simulation and Visual Analysis of Crowds (pp. 347-382). Springer, New York, NY.

Mehran, R., Oyama, A., & Shah, M. (2009, June). Abnormal crowd behavior detection using social force model. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 935-942). IEEE.

Pelechano, N., & Malkawi, A. (2008). Evacuation simulation models: Challenges in modeling high rise building evacuation with cellular automata approaches. Automation in construction, 17(4), 377-385.

Perdikaki, O., Kesavan, S., & Swaminathan, J. M. (2012). Effect of traffic on sales and conversion rates of retail stores. Manufacturing & Service Operations Management, 14(1), 145-162.

Sabokrou, M., Fayyaz, M., Fathy, M., Moayed, Z., & Klette, R. (2018). Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. Computer Vision and Image Understanding.

Soomaroo, L., & Murray, V. (2012, January 31). Disasters at Mass Gatherings: Lessons from History. Retrieved September 15, 2018, from http://currents.plos.org/disasters/article/disasters-at-mass-gatherings

Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., Alghraibeh, A. M., Aryeetey, R., Bertoni, A., Bettache, K., Blumen, S., Błażejewska, M., Bortolini, T., Butovskaya, M., Castro, F. N., Cetinkaya, H., Cunha, D., David, D., David, O. A., Dileym, F. A., Domínguez Espinosa, A. d. C., Donato, S., Dronova, D., Dural, S., Fialová, J., Fisher, M., Gulbetekin, E., Hamamcıoğlu Akkaya, A., Hromatko, I., Iafrate, R., Iesyp, M., James, B., Jaranovic, J., Jiang, F., Kimamo, C. O., Kjelvik, G., Koç, F., Laar, A., de Araújo Lopes, F., Macbeth, G., Marcano, N. M., Martinez, R., Mesko, N., Molodovskaya, N., Moradi, K., Motahari, Z., Mühlhauser, A., Natividade, J. C., Ntayi, J., Oberzaucher, E., Ojedokun, O., Omar-Fauzee, M. S. B., Onyishi, I. E., Paluszak, A., Portugal, A., Razumiejczyk, E., Realo, A., Relvas, A. P., Rivas, M., Rizwan, M., Salkičević, S., Sarmány-Schuller, I., Schmehl, S., Senyk, O., Sinding, C., Stamkou, E., Stoyanova, S., Šukolová, D., Sutresna, N., Tadinac, M., Teras, A., Tinoco Ponciano, E. L., Tripathi, R., Tripathi, N., Tripathi, M., Uhryn, O., Yamamoto, M. E., Yoo, G., and Pierce, J. D. 2017. "Preferred Interpersonal Distances: A Global Comparison," Journal of Cross-Cultural Psychology (48:4), pp. 577-592.

Transit Safety Data by Mode for All Reported Incidents. (2018). Retrieved September 15, 2018, from https://www.bts.gov/content/transit-safety-data-modea-all-reported-incidentsb

Treuille, A., Cooper, S., & Popović, Z. (2006). Continuum crowds. ACM Transactions on Graphics (TOG), 25(3), 1160-1168.

Trottier, L., Gigu, P., & Chaib-draa, B. (2017, December). Parametric exponential linear unit for deep convolutional neural networks. In Machine Learning and Applications (ICMLA), 2017 16th IEEE International Conference on (pp. 207-214). IEEE.

Verma, G. K., & Dhillon, A. (2017, November). A Handheld Gun Detection using Faster R-CNN Deep Learning. In Proceedings of the 7th International Conference on Computer and Communication Technology (pp. 84-88). ACM.

Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853.

Zhang, C., Li, H., Wang, X., & Yang, X. (2015). Cross-scene crowd counting via deep convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 833-841).

Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2016). Single-image crowd counting via multi-column convolutional neural network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 589-597).