COVID-GAN: Estimating Human Mobility Responses to COVID-19 Pandemic through Spatio-Temporal Conditional Generative Adversarial Networks

Han Bao
han-bao@uiowa.edu
University of Iowa

Xun Zhou
xun-zhou@uiowa.edu
University of Iowa

Yingxue Zhang
yzhang31@wpi.edu
Worcester Polytechnic Institute

Yanhua Li
yli15@wpi.edu
Worcester Polytechnic Institute

Yiqun Xie
xie@umd.edu
University of Maryland

ABSTRACT

The COVID-19 pandemic has posed grand challenges to policy makers, raising major social conflicts between public health and economic resilience. Policies such as closure or reopen of businesses are made based on scientific projections of infection risks obtained from infection dynamics models. While most parameters in infection dynamics models can be set using domain knowledge of COVID-19, a key parameter – *human mobility* – is often challenging to estimate due to complex social contexts and limited training data under escalating COVID-19 conditions. To address these challenges, we formulate the problem as a spatio-temporal data generation problem and propose COVID-GAN, a spatio-temporal Conditional Generative Adversarial Network, to estimate mobility (e.g., changes in POI visits) under various real-world conditions (e.g., COVID-19 severity, local policy interventions) integrated from multiple data sources. We also introduce a domain-constraint correction layer in the generator of COVID-GAN to reduce the difficulty of learning. Experiments using urban mobility data derived from cell phone records and census data show that COVID-GAN can well approximate real-world human mobility responses, and that the proposed domain-constraint based correction can greatly improve solution quality.

CCS CONCEPTS

• **Information systems** → **Spatial-temporal systems**; • **Computing methodologies** → **Machine learning**.

KEYWORDS

Mobility estimation, Conditional Generative Adversarial Networks, COVID-19

SIGSPATIAL '20, November 3–6, 2020, Seattle, WA, USA

ACM ISBN 978-1-4503-8019-5/20/11…$15.00
https://doi.org/10.1145/3397536.3422261

ACM Reference Format:
Han Bao, Xun Zhou, Yingxue Zhang, Yanhua Li, and Yiqun Xie. 2020. COVID-GAN: Estimating Human Mobility Responses to COVID-19 Pandemic through Spatio-Temporal Conditional Generative Adversarial Networks. In *28th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '20), November 3–6, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3397536.3422261

1 INTRODUCTION

The COVID-19 pandemic has spread to over 200 countries in the world, with 8,525,042 confirmed cases and 456,973 deaths as of June 20^{th}, 2020. Fig. 1 shows the latest situation report map from the World Health Organization [3]. The rapid spread of the infectious disease has posed grand challenges to policy makers due to the raising social conflicts between (1) the need to reduce new infections and protect public health, and (2) the demand of reopening to avoid breakdown of economics and support essential needs in daily lives. As more policy makers start exploring mitigation of this dilemma with staged reopening, a core mission is to avoid major resurgence of infections caused by eased social distancing policies (e.g., stay-at-home orders, limit of group gatherings, limit on restaurant capacity).

Currently, COVID-19 related policies are often informed by scientific projections of infection risks obtained from COVID-19 transmission dynamics models. While many parameters in infection dynamics models can be set using domain knowledge of COVID-19, a key parameter – *human mobility responses*[1] – is often challenging to estimate due to complex and sometimes unknown social contexts, as well as limited training data under escalating COVID-19 conditions.

Research Goal. Responding to the urgent need by policy makers and public health experts, we address the following *human mobility response estimation problem*: Given a set of inputs on contextual (e.g., population, POI counts), epidemic (e.g., COVID-19 cases) and policy (e.g., stay-at-home orders) conditions, we aim to estimate maps of human mobility responses by learning from existing ground truth data. Note, here the input conditions might not have been observed in the targeted area in the historical data.

Challenges. The problem has two major challenges. First, human mobility responses depend on many complex social-physical factors.

[1] In this work, human mobility responses are measured by the number of visits to point-of-interest (POI) such as grocery and hardware stores, restaurants, gas stations, etc.

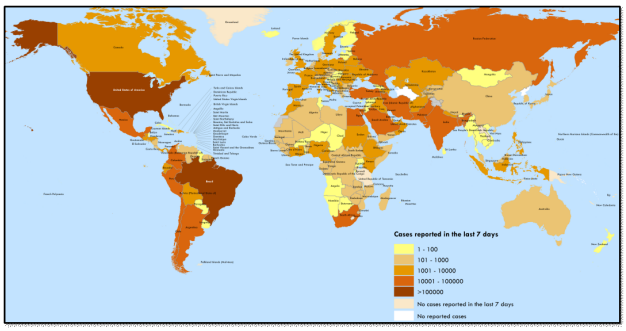


Figure 1: World map of COVID-19 cases by WHO [3]

While some factors are known and have related data available (e.g., population, POI distribution, policy), there exist many others (e.g., people's will or reaction over time, echo chamber effects, random events) that may be unknown, uncertain, random or have no data available. In addition, the contribution of many known and unknown underlying factors to human mobility responses may vary from place to place due to cultural differences, economic environments, and etc. Second, the COVID-19 pandemic is a rare, very recent and still-fast-changing event. As a result, the amount of related data is relatively small. This increases the difficulty of learning, especially when considering the random effects brought by the first challenge.

Due to the urgent need of effective control over the COVID-19 pandemic, there have been many works studying the importance of human mobility responses in COVID-19 transmission dynamics models as well as effectiveness of early-phase social distancing measures (e.g., isolation, stay-at-home order, lock-down) on containing the spread of the disease [5, 9, 17, 26]. There have also been studies evaluating the feasibility of scalable contact tracing methods in COVID-19 monitoring [7, 14]. While these studies have demonstrated the importance of human mobility responses, they do not address the main challenges in mobility estimation or simulation (e.g., effects of unknown, uncertain and random factors; limited training data). In terms of learning approaches, most regression based estimators (e.g., linear and non-linear regression, kernelized support-vector regressor) rely on a fixed set of features and cannot consider unknown and uncertain factors.

Contributions. To address these challenges, *we make the first attempt to formulate the COVID-19 human mobility responses estimation problem as a deep learning based data generation problem.* Specifically, we propose a COVID-GAN, a spatio-temporal conditional Generative Adversarial Network (GAN) [10, 12], to estimate mobility under various real-world conditions such as COVID-19 severity and local policy interventions. The use of a conditional GAN [10] allows consideration of unknown and uncertain factors (i.e., modeled as latent factors). In addition, we introduce a domain-constraint based correction layer in the generator to reduce the difficulty of learning and help mitigate the challenge of relatively small training data.

Our main contributions in this paper are as follows:

- We gather and integrate various types of data (e.g., contextual, epidemic and policy conditions) from multi-sources (e.g., SafeGraph [4], US Census Bureau [1], CDC [2], local government) to provide a multi-view input for the estimation problem.
- We design COVID-GAN, a spatio-temporal conditional generative adversarial network tailed for COVID-19 data, to learn and estimate human mobility responses. Specially, we introduce a new phase in the generator of COVID-GAN, which leverages known domain-knowledge to form low-computational-cost constraints to correct or mitigate spurious results during training. This domain constraint based correction can reduce the difficulty of learning and improve solution quality;
- We perform a variety of experiments on real-world COVID data to validate the solution quality improvements achieved by the proposed approach under different scenarios, and study the effects of related features on the quality of results.

Experiment results show that the proposed COVID-GAN can well mimic real-world human mobility responses in different settings, and that the proposed domain constraint based correction layer can greatly improve solution quality.

The rest of the paper is organized as follows. Section 2 defines the problem. Section 3 details the methodology and model structure. We present evaluation results in Section 4 and further discuss and highlight the technical insights of the proposed solution framework in Section 5. Related works are summarized in Section 6, and the paper is concluded in Section 7.

2 OVERVIEW

Mobility estimation in the COVID-19 pandemic is a complex task requiring a variety types of information such as population, COVID-19 situation, etc. Fortunately, many organizations have opened their data for research to help combat the pandemic. For example, SafeGraph [4], a leading provider of place and POI data, has started a COVID-19 Data Consortium to provide free access to their commercial data. Similarly, the Center for Disease Control and Prevention (CDC) also provides dynamic updates of COVID-19 related public heath statistics [2].

In the following part, we first introduce a set of basic concepts about our data modeling and then provide a formal problem definition.

2.1 Basic Concepts

Definition 2.1 (Spatial grid model $GRID_S$). A grid-discretization of a study area S, i.e., the geographic region of interest in an estimation task.

Definition 2.2 (Contextual conditions $\boldsymbol{C}_c$). Relatively long-term attributes that are not necessarily specific to COVID-19, including population, median income, timestamp, and number of POIs (e.g., grocery stores, schools, gas stations, restaurants).

Definition 2.3 (Epidemic conditions $\boldsymbol{C}_e$). COVID-19 statistics from COVID-19 situation reports published authoritative public health organizations (e.g., CDC), including number of confirmed cases, COVID-19 related deaths, etc.

Definition 2.4 (Policy conditions $\boldsymbol{C}_p$). Social distancing orders declared by officials during COVID-19 pandemic. Policies may play

a key role in mobility changes. For example, strict stay-at-home or shelter-in-place orders typically lead to a significant decrease in mobility.

*Definition 2.5 (Human mobility responses **M**).* In this work, M is measured by the number of visits to POIs (e.g., grocery stores, hardware stores, restaurants, gas stations), which according to our public health colleagues, is a major factor in COVID-19 transmission since individuals tend to have closer distances and interact with each other at POIs.[2] Human mobility response is target of the estimation task.

Definition 2.6 (Generator G). A learned process used to generate a map of human mobility responses $\mathbf{M}_G$ given a set of conditions.

Definition 2.7 (Discriminator D). Outputs a probability p_{real} that a map of human mobility responses is from real-world rather than a generator G.

2.2 Problem Statement

The problem is formally defined as follows:

Inputs:

- A spatial grid model $GRID_S$ for study area S;
- Contextual conditions $\mathbf{C}_c$ (e.g., population, median income, POI counts);
- Epidemic conditions $\mathbf{C}_e$ (COVID-19 statistics);
- Policy conditions $\mathbf{C}_p$ (e.g., social distancing policies);
- Human mobility responses $\mathbf{M}$;

Output:

- A generator G to estimate/generate maps (on $GRID_S$) of human mobility responses M_G under $\mathbf{C}_c$, $\mathbf{C}_e$ and $\mathbf{C}_p$;

Objective:

- Maximizing the probability p_{real} from D on M_G so that G approximates the underlying generation process of real data;

Constraints:

- Temporal resolution of estimation is aggregated at week-level;
- Mobility estimation at week i can only use COVID-19 statistics achieved before week i;
- Data is aggregated and does not contain private information.

In the scope of the present study, human mobility responses are measured by number of visits to point-of-interests (POIs), and other related measures are currently out of scope and will be evaluated in future work.

In addition, the current temporal resolution of our analysis is set to week-level due to the very high variance/noise of mobility from day to day within a week. Thus, to put more confidence on estimation results, the scope in this paper is to perform mobility estimation at an aggregated week-level as defined by the first constraint. The second constraint is added because changes in COVID-19 statistics (e.g., cases, death) can be considered as a direct or indirect result of mobility changes. Thus, in this problem, we make such post-information not accessible to mobility estimation for the same timestamp. However, this does not limit the use of COVID-19 statistics from previous timestamps, which may truly affect people's mobility responses.

[2]In comparison, overlapping trajectories of private vehicles or passengers on private vehicles may raise less concerns in the spread process.

3 COVID-GAN FOR MOBILITY ESTIMATION

Our approach – COVID-GAN – is inspired by the conditional Generative Neural Network (cGAN). cGAN is a natural modeling structure for human mobility response estimation in this COVID-19 scenario because in the problem there exist both a number of known (e.g., COVID-19 situation, policies) and unknown/random underlying factors (e.g., wills of individuals). With the cGAN structure, the known factors can be fed in as learning conditions and unknown/random variables can be represented by vectors of latent code.

However, the basic cGAN structure is insufficient due to complexity of the phenomenon as well as relatively small training data considering the recentness of this ongoing COVID-19 pandemic and the potentially high spatial variability across regions (e.g., different social, economical and COVID-19 contexts within a country or across many). To mitigate this challenge, we feature a new design to advance the original cGAN structure, i.e., a new phase in the conditional generator to constrain estimated values using domain knowledge before the results flow into the discriminator.

In the following part, we first introduce the integration and preprocessing of multi-sourced data in Sec. 3.1. Then, we introduce the network design and a domain-knowledge based correction scheme in Sec. 3.2 and followed by training step in Sec. 3.3, and finally show estimation step for real-data in Sec. 3.4.

3.1 COVID-GAN Data Preparation

Fig. 2 shows a summary of the multi-view data we gathered from multiple sources based on definitions in Sec. 2.1 (i.e., contextual, epidemic and policy conditions; human mobility responses). As we can see, most of the data is associated with different geographic units (e.g. census block groups, counties) due to different data sources or privacy protection concerns. Thus, further spatio-temporal data processing are needed before training and estimation.

First, to integrate all the data of various types and geographic units into the same format, we adopt a commonly used space-partitioning approach [16, 21, 39] to segment the input spatial domain into grid cells of size $1km \times 1km$, and all input data in Fig. 2 are then segmented based on these grid cells. Note that some features are re-scaled during this process. For example, population is linearly re-scaled using the corresponding area ratios between the area of the original polygon and the grid cells.

Since the temporal resolution of mobility estimation is week-level as discussed in problem definition (Sec. 2), we further merge all daily information (e.g., mobility, number of cases and death related to COVID-19) into weekly aggregations for all grid cells.

Finally, we select spatial unit windows consisting of $s \times s$ grid cells with their corresponding features (i.e., conditions and constraints for COVID-GAN) to train our COVID-GAN. In this work, we set the window size to 10×10 grid cells.

3.2 COVID-GAN Architecture

Fig. 3 shows the adversarial structure of the proposed COVID-GAN. COVID-GAN is composed of a conditional generator G and a discriminator D. To help improve the estimation quality, we propose

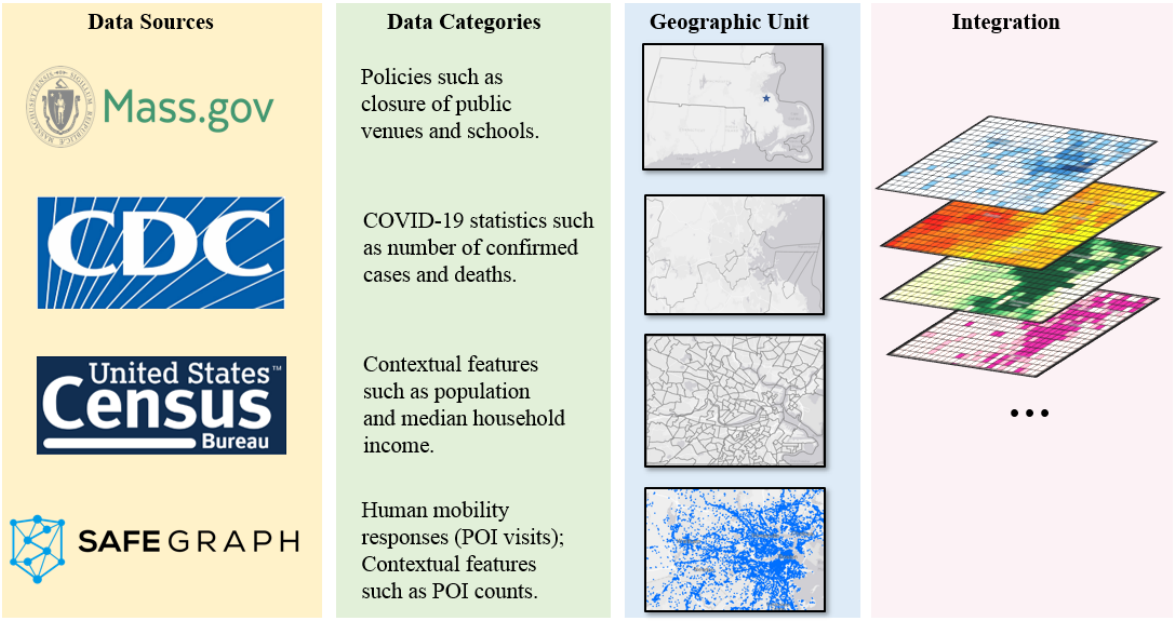


Figure 2: Multi-source data for COVID-GAN.

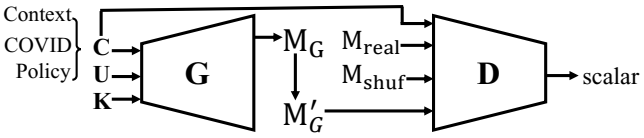


Figure 3: COVID-GAN

a new phase in the generator to constrain the estimated values using domain knowledge. Next, we will discuss the three key components in our COVID-GAN, i.e., the conditional generator, discriminator and domain-knowledge constraint.

3.2.1 Generator Phase-1: Conditions and Latent Factors.
As shown in Fig. 3, the input of conditional generator (skipping the dimension of "batch" for simplicity of description) includes a condition tensor $\mathbf{C} \in \mathbb{R}^{s\times s\times c_0}$, latent code tensor $\mathbf{U} \in \mathbb{R}^{s\times s\times u_0}$ and a domain knowledge tensor $\mathbf{K} \in \mathbb{R}^{s\times s\times k}$, where s is the side length of each spatial unit window introduced in Sec. 3.1, c_0 is the number of conditions (including contextual features, COVID statistics and policy conditions), u_0 is the dimension of the latent code vector used to model the randomness in human mobility responses and k is the number of layers in domain constraint.

In generator, the **local-feature-projection (LFP) layer** is used to map existing features in the current layer $L_i$to those of the next layer L_{i+1}. The local-feature-projection layer is a special case of a convolutional layer in which the kernel size is $1\times1\times F_i\times F_{i+1}$, where F_i and F_{i+1} are the number of features/channels in layer L_i and $Li + 1$, respectively. [3] We use local feature projection layers instead of convolutional layers with larger kernel size mainly to reduce the boundary effects on grid cells of each spatial window. This is important in mobility estimations since the final mobility map is typically an integration of many overlapping spatial windows (Sec.3.4), which may lead to accumulated boundary effects and thus reduce generation quality.

The detailed architecture of the conditional generator is shown in Fig. 4. In generator, the input tensors **C** and **U** are concatenated together and pass 4 local-feature-projection layers. The first three local-feature-projection layers are activated by Rectified Linear Unit (ReLU) and batch normalized. The last local-feature-projection layer is activated by hyperbolic tangent function, and we get an $s \times s$ matrix $\mathbf{M}_G$, where each entry represents the estimated human mobility of the corresponding grid cell. Then, domain-knowledge tensor **K** are enforced on $\mathbf{M}_G$ before the mobility values are fed into the discriminator to correct generated values and improve mobility estimation quality.

3.2.2 Generator Phase-2: Domain Knowledge Constraint.
Recently, domain knowledge assisted learning has become an emerging trend, which can help reduce the difficulty of learning and overfitting with a small amount of training data and shows encouraging improvements in learning results. The domain knowledge ranges from observations (e.g., lake temperature vs. water depth [15], morphological characteristics of objects [30]) to social and physical theories (e.g., routine activity theory [8], engine combustion models [20]).

In COVID-GAN, we add a $s \times s \times k$ tensor K to represent domain knowledge that can be potentially used to correct or improve mobility estimations in phase-1. A domain constraint should satisfy the following requirements: (1) it contains finite layers (e.g., in our study, the layers of the domain constraint **K** corresponds to the last dimension of tensor **K**); (2) it is related to the input conditions

[3]This type of layers has also been used in other networks (e.g., YOLO [25, 31]) for feature construction.

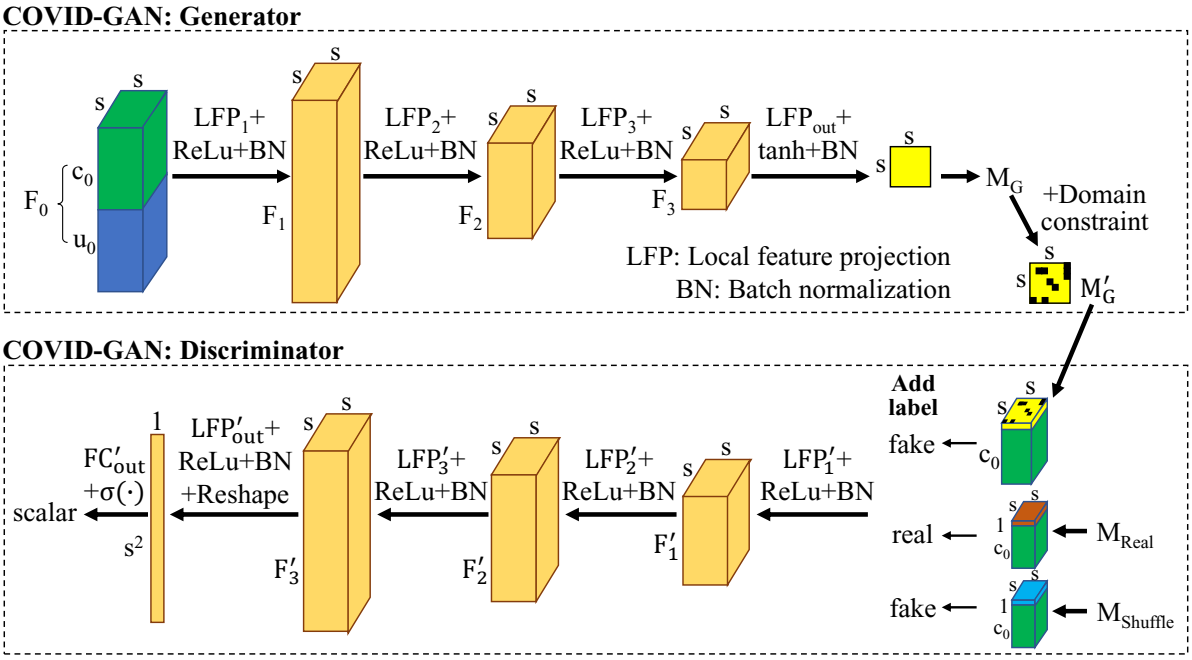


Figure 4: Network Architecture of COVID-GAN.

C so that the usability of COVID-GAN does not depend on the constraint; and (3) for calculation simplicity, the constraint should be enforced or applied with a low-cost operation (e.g., a simple element-wise tensor calculation) using $\mathbf{M}_G$ and K. The third constraint is especially important in this context of deep learning since the constraint enforcement operation will repeat in every iteration.

In this study, we construct a constraint **K** with $k = 1$ satisfying all the three requirements mentioned above using a combination of time, POI and policy (TPP constraint).

Denote m_{xy} as mobility values at cell (x, y) in $\mathbf{M}_G$. We have $m_{xy} \in (-1, 1)$ as it is an output of $\tanh(\cdot)$ in the output layer. Since in COVID-GAN, human mobility at a grid cell is represented by the number of visits to POIs in a certain time slot (e.g., one week), m_{xy} equals to -1 when there is no visit in a cell, $m_{xy} > -1$ if (1) there exists POIs in the grid cell during the current time slot; and (2) at least one of the POIs is allowed to open. To construct the corresponding TPP constraint **K** with $k = 1$, we treat **K** as an $s \times s$ matrix where each entry value equals to 1 if condition (1) and (2) mentioned above are both satisfied; otherwise, the entry value is -1. Fig. 5 shows an illustrative example of our TPP constraint.

The constraint can be then enforced by:

$$\mathbf{M}'_G = \min(\mathbf{M}_G, \mathbf{K}) \tag{1}$$

By applying Eq. (1), the mobility values in $\mathbf{M}_G$ remain the same in grid cells if its corresponding TTP value is 1, or the values in $\mathbf{M}_G$ are corrected to -1. Thus, this constraint overall only requires one data layer in **K** and can be easily enforced by a single $\min(\cdot)$ function. Note that the actual values (e.g., -1, 1) in the constraint may vary for different choices of activation functions in the output layer. Here the values are chosen for $\tanh(\cdot)$. This type of constraint

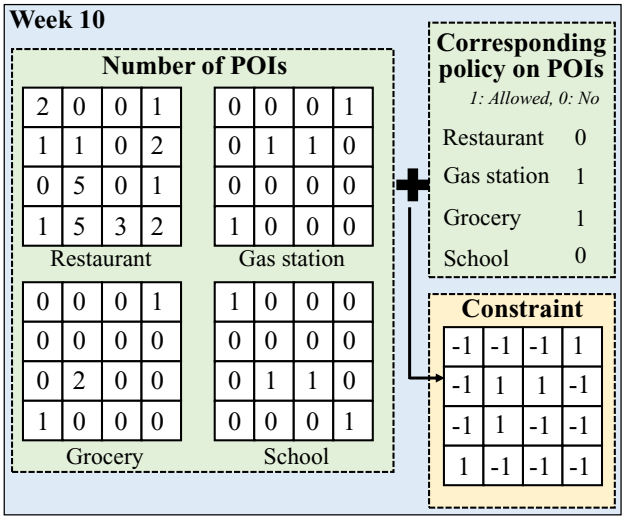


Figure 5: An illustrative example of the TPP constraint.

corrections can potentially help the generator focus on valid cells and improve the generation quality.

3.2.3 *Discriminator.*

Fig .4 shows the detailed architecture of the discriminator. Its input is a tensor of size $s \times s \times (c_0+1)$, where c_0 is the number of conditions (same as generator).

As shown in Fig. 4, the input tensor can be created in three different ways:

– Conditions concatenated with generated mobility M'_G;

- Conditions concatenated with their corresponding real mobility M_{real};
- Conditions concatenated with mismatched real mobility M_{shuffle}.

Among the three kinds of inputs, only the second one is labeled with "real" whereas the other two are marked with "fake". The input tensor is then fed into the discriminator, and the goal of the discriminator is to learn to tell if the input is "real" or "fake".

The discriminator also contains four local-feature-projection layers, they are activated by ReLU and followed by batch normalizations, the output of the last local-feature-projection layer is reshaped into a vector and passes a fully connected layer as well as a Sigmoid function and outputs a final scalar, which indicates the probability whether the input of discriminator is "real" data.

3.3 COVID-GAN Training

The training of COVID-GAN is performed via adversarial confrontation between the generator and discriminator. The objective function of COVID-GAN is then formulated as an antagonizing bi-level min-max optimization problem with binary-cross-entropy:

$$\begin{aligned}\min_G \max_D f(G,D) &= \mathrm{E}_{\mathbf{M}\sim P_{data}}[\log D(\mathbf{M},\mathbf{C})] \\ &+ \mathrm{E}_{\mathbf{U}\sim P_U}[\log(1-D(G(\mathbf{C},\mathbf{U},\mathbf{K}),\mathbf{M}))]\end{aligned} \tag{2}$$

Alg. 1 presents the training process of COVID-GAN. The training of the discriminator uses the three types of $(\mathbf{M},\mathbf{C})$ combinations as illustrated in Sec. 3.2.3: $(\mathbf{M}'_G,\mathbf{C})$, $(\mathbf{M}_{\text{real}},\mathbf{C})$ and $(\mathbf{M}_{\text{shuffle}},\mathbf{C})$. Denote η_D as the learning rate of discriminator, $\boldsymbol{\theta}_D$ as the parameters of discriminator, the loss function and the update rule of D are shown in Eq. (3) and Eq. (4), respectively.

$$\begin{aligned}f_D = -\frac{1}{m}\sum_{i=1}^{m}\Big(&\log(1-D((\mathbf{M}'_G)^i,\mathbf{C}^i)) + \log(D(\mathbf{M}^i_{\text{real}},\mathbf{C}^i)) \\ &+ \log(1-D(\mathbf{M}^i_{\text{shuffle}},\mathbf{C}^i))\Big)\end{aligned} \tag{3}$$

$$\boldsymbol{\theta}_D = \boldsymbol{\theta}_D + \eta_D \nabla f_D(\boldsymbol{\theta}_D) \tag{4}$$

where m is the total number of samples in a batch, and index i refers to the i^{th} sample.

Similarly, denote η_G as the learning rate of the generator G, and $\boldsymbol{\theta}_G$ as the parameters in G, we have the loss function and update rule of G as:

$$\begin{aligned}f_G &= \frac{1}{m}\sum_{i=1}^{m}\Big(\log(1-D((\mathbf{M}'_G)^i,\mathbf{C}^i))\Big) \\ &= \frac{1}{m}\sum_{i=1}^{m}\Big(\log(1-D(G(\mathbf{C}_i,\mathbf{U}_i,\mathbf{K}_i),\mathbf{C}^i))\Big)\end{aligned} \tag{5}$$

$$\boldsymbol{\theta}_G = \boldsymbol{\theta}_G + \eta_G \nabla f_G(\boldsymbol{\theta}_G) \tag{6}$$

3.4 COVID-GAN Estimation

The goal of this step is to generate the mobility map of the entire study area (e.g., a city or county) from the estimations of $s \times s$ spatial unit windows. Since COVID-GAN would generate the mobility maps of spatial unit windows in different time slots and areas,

Algorithm 1: COVID-GAN Training

Require:
- List of conditions $\mathbf{C}$
- List of domain constraints $\mathbf{K}$
- List of real mobility $\mathbf{M}_{\text{real}}$
- Number of epochs *epoch*

```
G = initG()
D = initD()
for e = 1 to epoch do
    for batch in {C, K, M_real} do
        M'_G = G(batch.C, rand(P_U), batch.K)
        {# for discriminator:}
        batch_mis = mismatchShuffle(batch)
        Update D using Eqs. (3) and (4)
        {# for generator:}
        Update G using Eqs. (5) and (6)
    end for
end for
```

we present two schemes to generate the final map of mobility estimation.

Single-draw based sliding window: In this scheme, we move a $s \times s$ sliding window across the whole target area to prepare all conditions and constraints for the generator G. Then, using these well-prepared data, we train the COVID-GAN and generate a single $s \times s$ mobility estimation result for each unit window (i.e., a single draw from the distribution with latent factors), and build the whole map with all the generated results for unit sliding windows. Since each grid cell in the target area may have multiple generated results due to overlaps among the sliding unit windows, the final mobility value in a grid cell is the average of the generated results.

Multiple-draw based sliding window: This scheme takes w draws of the same window from the generator G instead of a single draw. The results of the multiple draws are averaged before being integrated into the estimation of the original study area.

In our experiments we use the multiple-draw version to reduce random effects in comparison. The single-draw version is better suited for scenarios when a decision maker would like to explore potential variations in human mobility responses.

4 EVALUATION

The overall evaluation framework is shown in Fig. 6. Our experiments aim to answer the following questions:

- How does COVID-GAN performs compared to the baseline in terms of solution quality?
- What is the effect of training on spatially-seen vs. spatially-unseen samples?
- What is the effect of training on temporally-seen vs. temporally-unseen samples?
- What are the effects of features in COVID-GAN on the performance?

The second question aims to compare the quality of generated mobility responses between spatial regions that are seen and unseen in the training data (either temporally seen or not). Similarly, the

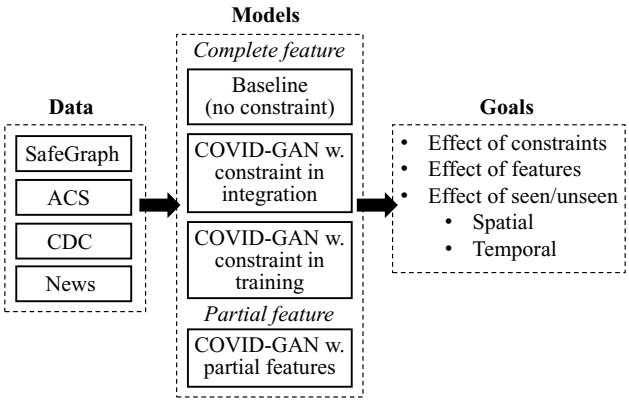
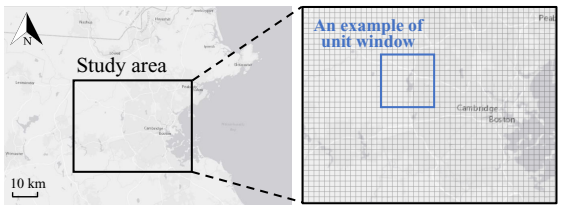


Figure 6: Overall evaluation framework.



(a) Study area around Boston (b) Grid partition (cell size: 1km)

Figure 7: Partitioning of Boston study area.

third questions aims to evaluate the same effect along the temporal dimension.

4.1 Boston Dataset

The study area used in the experiments is shown in Fig. 7, where Fig. 7 (a) visualizes the geographic context of the area around Boston, MA, and Fig. 7 (b) shows the grid partitioning of the space. The dimension of the study area is $37km \times 48km$ and each grid cell has a size of $1km \times 1km$. An example of a 10×10 unit window (Sec. 3.1) is illustrated by the blue box in Fig. 7(b). The duration of data is from 03/02 to 05/24, covering 12 weeks in total. The study area has over 2 million population in total and 26054 POIs. During the 12 weeks, the total number of POI visits is over 6 million. A "closure-of-public-venues" policy was implemented starting from 03/23 and lasted till the end of the 12-week period.

4.2 Estimation Quality Evaluation

Here we aim to answer the first three questions summarized at the beginning of Sec. 4. Fig. 9 and 10 shows a comparison of results of different models. Specifically, the first column (i.e., (a1) and (a2)) in both figures shows the ground truth, and the second and third columns show the results of the proposed COVID-GAN as well as the baseline approach (i.e., a conditional GAN [10, 39] without domain constraint based correction). In Fig. 9, we highlight the results of the candidate methods for a sub-region in the study area (the same region is used for both temporally seen and unseen data). The results for the full Boston study area is shown in Fig. 8. The colors used in map symbologies are classified using quantiles extracted from the ground truth (i.e., 0^{th}, 25^{th}, 50^{th}, 75^{th} and 100^{th}), a typical approach for enhanced map visualization. To reduce random effects in the comparison, all the results are based on multiple-draw based approach (10 repetitive runs) described in 3.4.

Full-map comparison: Fig. 8 (a), (b) and (c) show the ground truth and results of candidate methods for the full Boston study area for the final week in the data (not used in training). As we can see, mobility distribution generated by the COVID-GAN can capture the details in the distribution of human mobility responses much better than the baseline method. Fig. 8(d) shows the total estimated mobility (i.e., total number of POI visits) for the full map computed using the ground truth, COVID-GAN and the baseline. The total mobility (number of visits) of the ground truth, COVID-GAN and baseline are 422,223, 402,471 and 364,284, respectively. The differences (Fig. 8(e)) between the ground truth and the methods are 19,752 and 57,939 for COVID-GAN and the baseline, respectively. We can see that the overall estimation of COVID-GAN is much closer to that of the ground truth. To evaluate the distribution-closeness between the mobility values estimated by the ground truth and the methods, we further compute the Kullback–Leibler divergence [18] and the results are shown in Fig. 8(f). The X-axis are the number of equal-size bins used to discretize the mobility estimations, which is needed for the computation. As we can see, COVID-GAN achieves much lower KL-divergence values compared to the baseline consistently for different numbers of bins.

COVID-GAN vs. baseline: As we can see in Fig. 9 and Fig. 10, results of COVID-GAN and in (b1), (b2) are able to better approximate the mobility distributions in the ground truth compared to the results of the baseline method in (c1) and (c2), where a large number of cells contain a spurious mobility value, ranging from low to high. In addition, the trend remains the same for both temporally seen and unseen data in Fig. 9 as well as spatially seen and unseen data in Fig. 10.

Temporally-seen vs. temporally-unseen: In this comparison, temporally-seen refers to the time periods (i.e., weeks) that exist in the training data and have been seen by COVID-GAN. In contrast, temporally unseen refers to data of weeks that are outside the training data. To make the test more realistic, here the timestamp of unseen data must also be strictly after all timestamps in the training data so that the model does not try to estimate the past based on the future. In this experiment, we use the first 11 weeks of data to train the candidate models and leave the 12^{th} week out as "temporally unseen" data. According to the results in Fig. 9, we can see that COVID-GAN is able to maintain a high solution quality when being used to estimate human mobility responses for the unseen week, which is especially important in assisting policy-making during the ongoing pandemic.

Spatially-seen vs. spatially-unseen: To generate spatially unseen regions, we crop a $20km \times 20mk$ sub-region (i.e., a 20×20 sub-grid) off the total geographic space shown in Fig. 7, and data in this sub-region are not seen by COVID-GAN during training. This eliminates about one third of the total amount of training samples (overlaps with the sub-grid are not allowed). Fig. 10 shows the comparison of results by COVID-GAN and the baseline approach. Comparing results in the two rows, COVID-GAN achieves a better estimation for the first row when the data is spatially seen, and the

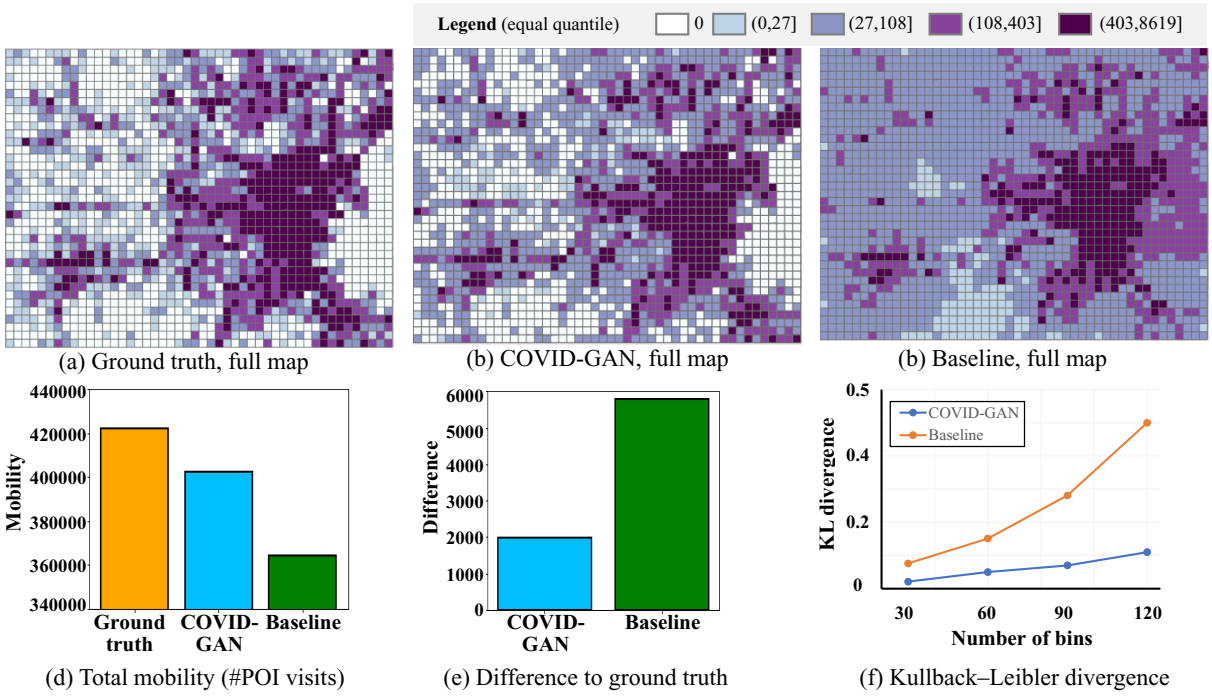


Figure 8: Mobility estimation results of the full Boston study area.

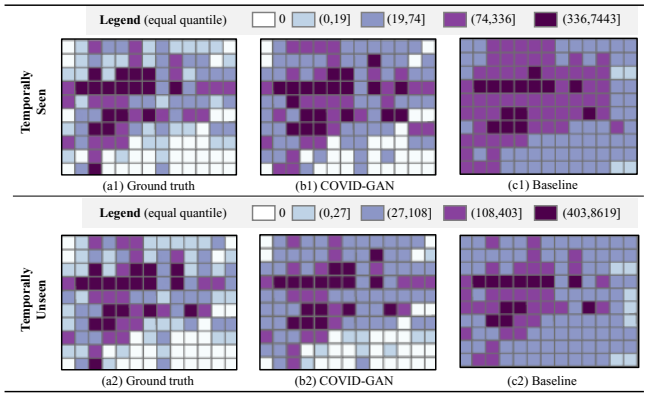


Figure 9: Detailed mobility estimation results in a sub-region of the study area for both temporally seen and unseen data.

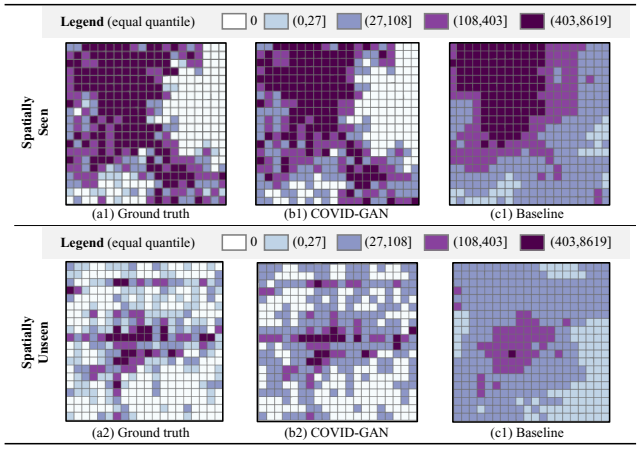


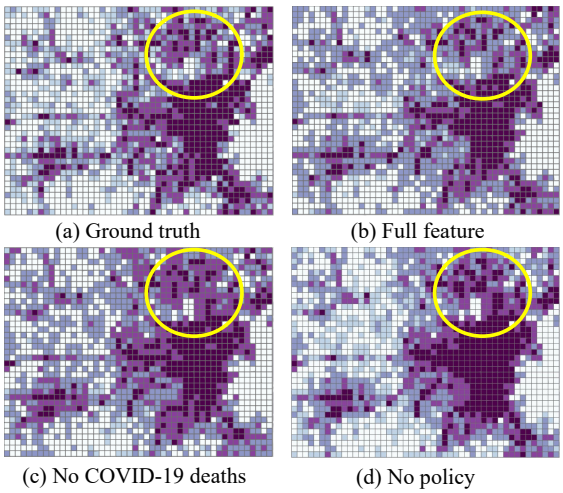
Figure 10: Mobility estimation results in spatially seen (first row) and unseen regions (second row) of the study area.

results of the baseline cannot well approximate the details in the mobility distribution.

4.3 Effect of Features

Here we evaluate the effect of two features, i.e., COVID-19 related deaths and policy, on the estimation results. Specifically, for each of two features, we trained a separate instance that does not contain it as a condition in **C**. Fig. 11 shows the results of COVID-GAN in (b) and the versions trained without COVID-19 related deaths and policy in (c) and (d), respectively. An example of the differences is highlighted by the yellow circle.

First, we can see the full feature based version can well approximate the mobility distribution in the ground truth. For the model trained without the condition on COVID-19 related deaths, we can see that it tends to over-estimate the mobility values in all three circles. The reason may be that COVID-19 related deaths has a discouraging effect on people's will to get out and visit POIs such as restaurants. As a result, removal of this feature may potentially



(a) Ground truth (b) Full feature

(c) No COVID-19 deaths (d) No policy

Figure 11: Effect of features on mobility estimation.

lead to less suppressing force on the value and an increase in the estimation.

A similar trend can be seen for the model trained without the policy condition. For example, in the highlighted circle, the "no-policy" results are over-estimations compared to the ground truth and the full-feature version. This result also conforms to our expectation since social distancing policies in our data are restricting factors on human mobility responses [22], and its absence may potentially increase the chance of over-estimation.

5 DISCUSSION

In this section, we discuss a few data and technical aspects of COVID-GAN in a more holistic manner.

Generation vs. prediction: As a conditional generative neural network, COVID-GAN is more of a generator than predictor by definition since the generation process involves randomly sampled latent factors to model unknown or uncertain factors in human mobility responses. In other words, the generated mobility results may not necessarily remain the same under the same condition (results in our experiments are averaged from multiple simulations). This allows policy makers or epidemiologists to evaluate different possibilities of human mobility responses during decision-making (e.g., change of policy). In contrast, a prediction model typically produces the same result for the same data once trained, and it is better suited when all or most of factors of human mobility responses are well-understood, stable and available in data.

Spatial variability: While our experiment results show that COVID-GAN performed reasonably well for a spatially-unseen region, that region is relatively small and is still adjacent to the spatially-seen data. Thus, more research is needed to examine COVID-GAN's performance for regions that are more distant to each other (e.g., different states, different countries). Given the typical social and physical differences in different geographic areas, new approaches may be needed to explicitly handle the spatial variability at larger geographic scales [13].

Data quality: In this work we did not investigate the effect of data quality (e.g., missing data, inaccuracy, anomaly) of input conditions, which may have bigger impacts in rural areas where the total volume of data tend to be smaller. This needs to be further investigated in future research to improve the generality of COVID-GAN.

6 RELATED WORK

COVID-19. There have been many studies [5, 6, 9, 17, 26] exploring the interplay between human mobility responses, social distancing policies, and transmission dynamics in response to the COVID-19 pandemic. For example, it was shown by [17] that strict implementation of social distancing policies can reduce mobility and substantially mitigate the spread of COVID-19. A US mobility change map was created in [9] to increase risk awareness of the public and visualize dynamic changes in mobility as COVID-19 situation and policy evolves. Studies [7, 14] have also explored the feasibility of utilizing contact tracing to control the spread of the disease through simulated synthetic data and real-world smartphone trajectories. These studies are timely in showing the important role played by mobility in the spread of COVID-19, but they do not address the challenges in real-world mobility estimation/simulation (e.g., effects of unknown, uncertain and random factors) and have not explored the potential use of deep learning based generative models to assist the estimation.

Deep Learning for Spatio-Temporal Prediction. There have been many deep learning based techniques developed for spatiotemporal data, including traffic accident prediction [36], flow prediction [23, 37], geospatial object mapping [30–32, 35], taxi driver behavior imitation [38], taxi demand [11, 28, 34], travel time estimation [29], dispersal event forecasting [27], etc. Most of these spatio-temporal deep learning techniques typically are stationary predictors (i.e., same result from two runs on same data) rather than generative models, or do not leverage domain knowledge based constraints to assist learning (e.g., cGAN [10, 12, 39]). In addition, generative neural networks have not been explored to assist human mobility response estimation at fine scale in this COVID-19 pandemic.

7 CONCLUSIONS AND FUTURE WORK

We proposed a conditional COVID-GAN to estimate/generate real-world human mobility responses to assist policy making during staged reopening in face of the COVID-19 pandemic. COVID-GAN integrated a variety of features including contextual features, COVID-19 statistics and policies from multi-sources such as SafeGraph, American Community Survey, Center for Disease Control and Prevention, news, etc. In addition, the generator incorporated a domain-constraint correction layer to efficiently reduce spurious results as well as the difficulty of learning. Experiment results showed that COVID-GAN can well mimic real-world human mobility responses and the domain-constraint based correction can greatly improve solution quality.

In future work, we aim to first explore statistically robust formulations (e.g., [8, 19, 33]) to improve the confidence of the estimation. We also plan to explore spatial variability aware formulations of COVID-GAN as well as other measures of human mobility response and efficiently enforceable domain constraints with epidemiologists.

Furthermore, the current work limits on single city. We will include more cities to improve the model capacity in transferring between different places, and consider distributed frameworks [24] for better scalability during data integration and model training. Finally, we will add more baselines (e.g., more traditional learning models) as well as metrics for model evaluation (e.g., RMSE, summary statistics across bins).

ACKNOWLEDGMENTS

This paper is funded partially by Safety Research using Simulation University Transportation Center (SAFER-SIM). SAFER-SIM is funded by a grant from the U.S. Department of Transportation's University Transportation Centers Program (69A3551747131). However, the U.S. Government assumes no liability for the contents or use thereof. Yingxue Zhang and Yanhua Li were supported in part by NSF grants IIS-1942680 (CAREER), CNS-1952085, CMMI-1831140, and DGE-2021871. We thank SafeGraph Inc. (www.safegraph.com) for providing free access to Boston POI data, including Core Places, Geometry, and Weekly Places Patterns for this research.

REFERENCES

[1] 2020. American Community Survey (ACS). https://www.census.gov/programs-surveys/acs.
[2] 2020. Centers for Disease Control and Prevention. https://www.cdc.gov/.
[3] 2020. COVID-19 situation report by WHO. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.
[4] 2020. SafeGraph. https://www.safegraph.com/.
[5] Patrick Bryant and Arne Elofsson. 2020. Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries. *medRxiv* (2020).
[6] Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al. 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368, 6489 (2020), 395–400.
[7] Hyunghoon Cho, Daphne Ippolito, and Yun William Yu. 2020. Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs. *arXiv preprint arXiv:2003.11511* (2020).
[8] Emre Eftelioglu, Shashi Shekhar, Dev Oliver, Xun Zhou, et al. 2014. Ring-shaped hotspot detection: a summary of results. In *2014 IEEE International Conference on Data Mining*. IEEE, 815–820.
[9] Song Gao, Jinmeng Rao, Yuhao Kang, Yunlei Liang, and Jake Kruse. 2020. Mapping county-level mobility pattern changes in the United States in response to COVID-19. *SIGSPATIAL Special* 12, 1 (2020), 16–26.
[10] Jon Gauthier. 2014. Conditional generative adversarial nets for convolutional face generation. *Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, Winter semester* 2014, 5 (2014), 2.
[11] Xu Geng, Yaguang Li, Leye Wang, Lingyu Zhang, Qiang Yang, Jieping Ye, and Yan Liu. 2019. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 3656–3663.
[12] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.
[13] Jayant Gupta, Yiqun Xie, and Shashi Shekhar. 2020. Towards Spatial Variability Aware Deep Neural Networks (SVANN): A Summary of Results. In *1st ACM SIGKDD Workshop on Deep Learning for Spatiotemporal Data, Applications, and Systems*.
[14] Joel Hellewell, Sam Abbott, Amy Gimma, Nikos I Bosse, Christopher I Jarvis, Timothy W Russell, James D Munday, Adam J Kucharski, W John Edmunds, Fiona Sun, et al. 2020. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* (2020).
[15] Anuj Karpatne, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar. 2017. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on knowledge and data engineering* 29, 10 (2017), 2318–2331.
[16] Amin Vahedian Khezerlou, Xun Zhou, Lufan Li, Zubair Shafiq, Alex X Liu, and Fan Zhang. 2017. A traffic flow approach to early detection of gathering events: Comprehensive results. *ACM Transactions on Intelligent Systems and Technology (TIST)* 8, 6 (2017), 1–24.
[17] Moritz UG Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M Pigott, Louis Du Plessis, Nuno R Faria, Ruoran Li, William P Hanage, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368, 6490 (2020), 493–497.
[18] Solomon Kullback. 1997. *Information theory and statistics*. Courier Corporation.
[19] Martin Kulldorff. 1997. A spatial scan statistic. *Communications in Statistics-Theory and methods* 26, 6 (1997), 1481–1496.
[20] Yan Li, Pratik Kotwal, Pengyue Wang, et al. 2020. Physics-guided energy-efficient path selection using on-board diagnostics data. *ACM Transactions on Data Science* (2020).
[21] Yanhua Li, Jun Luo, Chi-Yin Chow, Kam-Lam Chan, Ye Ding, and Fan Zhang. 2015. Growing the charging station network for electric vehicles with trajectory data analytics. In *2015 IEEE 31st International Conference on Data Engineering*. IEEE, 1376–1387.
[22] McKinsey. 2020. COVID-19's effect on jobs at small businesses in the United States. https://www.mckinsey.com/industries/social-sector/our-insights/covid-19s-effect-on-jobs-at-small-businesses-in-the-united-states#.
[23] Zheyi Pan, Zhaoyuan Wang, Weifeng Wang, Yong Yu, Junbo Zhang, and Yu Zheng. 2019. Matrix factorization for spatio-temporal neural networks with applications to urban flow prediction. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 2683–2691.
[24] Sushil K Prasad, Danial Aghajarian, Michael McDermott, et al. 2017. Parallel processing over spatial-temporal datasets from geo, bio, climate and social science communities: A research roadmap. In *2017 IEEE International Congress on Big Data (BigData Congress)*. IEEE, 232–250.
[25] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 779–788.
[26] Jean-Paul R Soucy, Shelby L Sturrock, Isha Berry, Nick Daneman, Derek R MacFadden, and Kevin A Brown. 2020. Estimating the effect of physical distancing on the COVID-19 pandemic using an urban mobility index. *medRxiv* (2020).
[27] Amin Vahedian, Xun Zhou, Ling Tong, W Nick Street, and Yanhua Li. 2019. Predicting urban dispersal events: A two-stage framework through deep survival analysis on mobility data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5199–5206.
[28] D. Wang, W. Cao, J. Li, and J. Ye. 2017. DeepSD: Supply-Demand Prediction for Online Car-Hailing Services Using Deep Neural Networks. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*. 243–254.
[29] Zheng Wang, Kun Fu, and Jieping Ye. 2018. Learning to Estimate the Travel Time. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery Data Mining* (London, United Kingdom) *(KDD '18)*. Association for Computing Machinery, New York, NY, USA, 858–866. https://doi.org/10.1145/3219819.3219900
[30] Yiqun Xie, Han Bao, Shashi Shekhar, and Joseph Knight. 2018. A TIMBER framework for mining urban tree inventories using remote sensing datasets. In *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1344–1349.
[31] Yiqun Xie, Rahul Bhojwani, Shashi Shekhar, and Joseph Knight. 2018. An unsupervised augmentation framework for deep learning based geospatial object detection: a summary of results. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 349–358.
[32] Yiqun Xie, Jiannan Cai, Rahul Bhojwani, Shashi Shekhar, and Joseph Knight. 2020. A locally-constrained yolo framework for detecting small and densely-distributed building footprints. *International Journal of Geographical Information Science* 34, 4 (2020), 777–801.
[33] Yiqun Xie and Shashi Shekhar. 2019. Significant DBSCAN towards Statistically Robust Clustering. In *Proceedings of the 16th International Symposium on Spatial and Temporal Databases*. 31–40.
[34] Huaxiu Yao, Fei Wu, Jintao Ke, Xianfeng Tang, Yitian Jia, Siyu Lu, Pinghua Gong, Jieping Ye, and Zhenhui Li. 2018. Deep multi-view spatial-temporal network for taxi demand prediction. *arXiv preprint arXiv:1802.08714* (2018).
[35] Jiangye Yuan. 2017. Learning building extraction in aerial scenes with convolutional networks. *IEEE transactions on pattern analysis and machine intelligence* 40, 11 (2017), 2793–2798.
[36] Zhuoning Yuan, Xun Zhou, and Tianbao Yang. 2018. Hetero-convlstm: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 984–992.
[37] Junbo Zhang, Yu Zheng, and Dekang Qi. 2017. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Thirty-First AAAI Conference on Artificial Intelligence*.
[38] Xin Zhang, Yanhua Li, Xun Zhou, and Jun Luo. 2019. Unveiling Taxi Drivers' Strategies via cGAIL: Conditional Generative Adversarial Imitation Learning. In *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1480–1485.
[39] Yingxue Zhang, Yanhua Li, Xun Zhou, Xiangnan Kong, and Jun Luo. 2019. TrafficGAN: Off-Deployment Traffic Estimation with Traffic Generative Adversarial Networks. In *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1474–1479.