Conditional Generative Models are Sufficient to Sample from Any Causal Effect Estimand

Md Musfiqur Rahman
Purdue University

Matt Jordan
University of Texas at Austin

Murat Kocaoglu
Purdue University

Causality in Machine Learning

- **Fairness**: Is there any **confounder** in your data which might give you **wrong prediction**?

Causality in Machine Learning

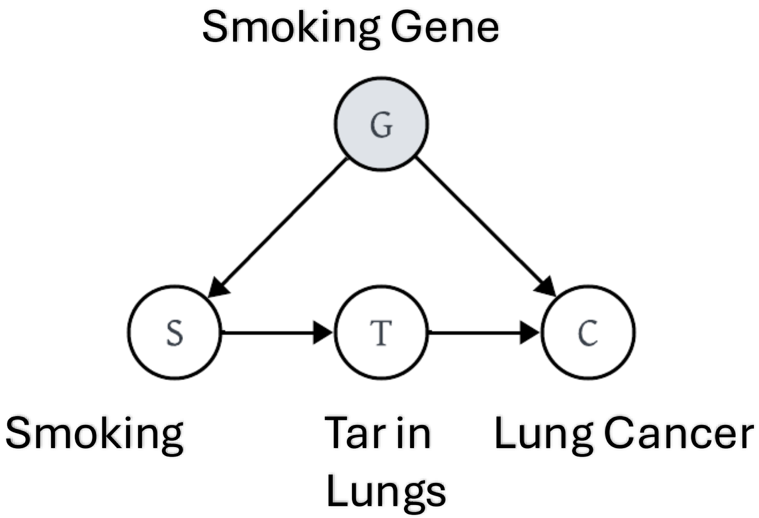
- **Fairness**: Is there any **confounder** in your data which might give you **wrong prediction**?
- **Robustness:** Is there any **bias in your data** that might affect your **model accuracy** in the **test domain**?

Causality in Machine Learning

- Confounder : wrong prediction?
- Bias in your data: worse model accuracy in the test domain?
- **Predictions based on cause effect are free from such harms.**

Causality in Machine Learning

- Confounder : wrong prediction?
- Bias in your data: worse model accuracy in the test domain?
- **Predictions based on cause effect are free from such harms.**

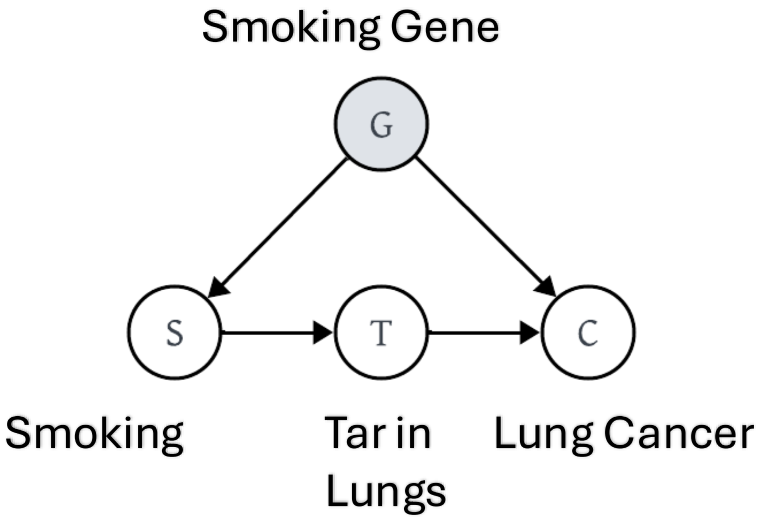


Causal effect of smoking on Lung cancer,

$$P(C|do(S)) = ?$$

Causality in Machine Learning

- Confounder : wrong prediction?
- Bias in your data: worse model accuracy in the test domain?
- **Predictions based on cause effect are free from such harms.**



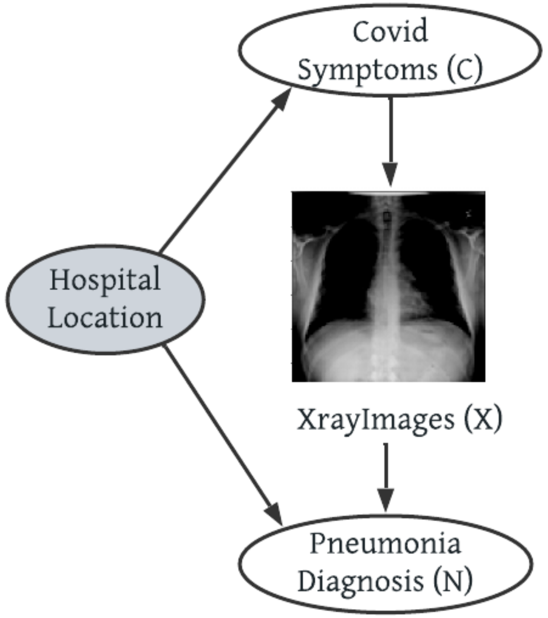
Causal effect of smoking on Lung cancer,

$$P(C|do(S)) = ?$$

Identification algorithms [1]

$$P(C|do(S)) = \sum_{T} P(T|S) \sum_{S'} P(C|S', T) P(S')$$

High dimensional interventional sampling

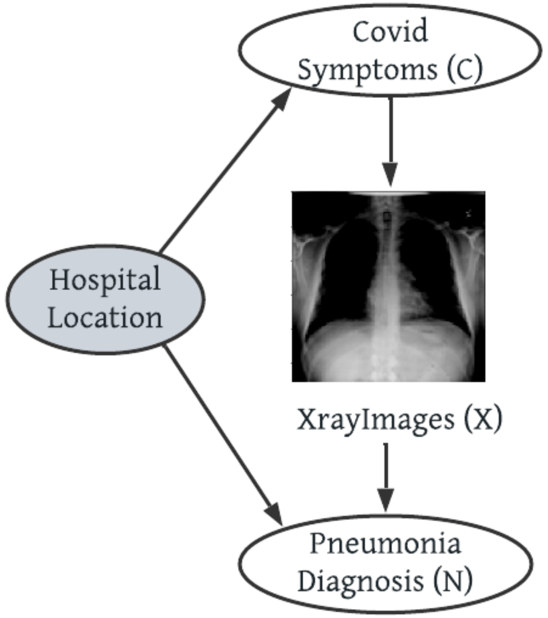


COVIDx CXR-3 dataset

$$P(N|do(C)) = \sum_X P(X|C) \sum_{C'} P(N|X, C')P(C')$$

High dimensional interventional sampling

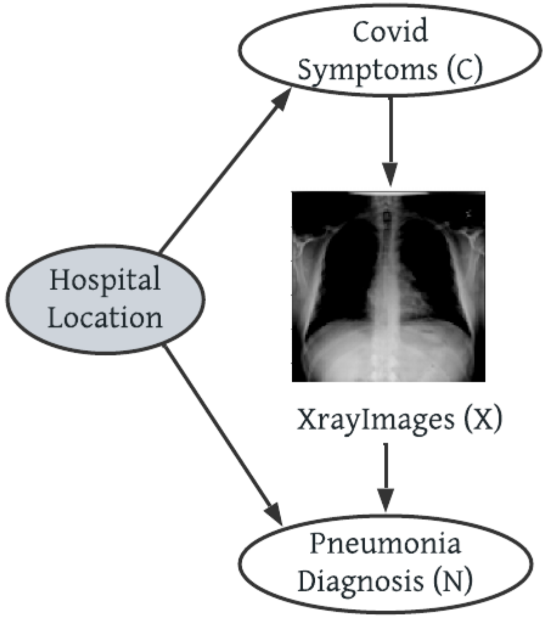
- P(X|C)= ? 🙁
- Train a conditional model G_X ? $X \sim G_X(C)$. 😎



COVIDx CXR-3 dataset

$$P(N|do(C)) = \sum_X P(X|C) \sum_{C'} P(N|X, C')P(C')$$

High dimensional interventional sampling



COVIDx CXR-3 dataset

- Conditional distributions:
- Interventional distributions: P(N|do(C)) or P(M|do(V))?

$$P(N|do(C)) = \sum_{X} P(X|C) \sum_{C'} P(N|X,C')P(C')$$

An approach to sample from high-dimensional interventional distribution!

Problem Definition

- Input:
  - Observational training data
  - A causal graph
- Goal:
  - Perform an intervention do(X=x)
  - Estimate numeric values of the causal effect.
  - Or Sample from the high-dimensional interventional distribution
  - .

$$do(X = x) \qquad P(y|do(x)) \qquad Y \sim P(y|do(x))$$

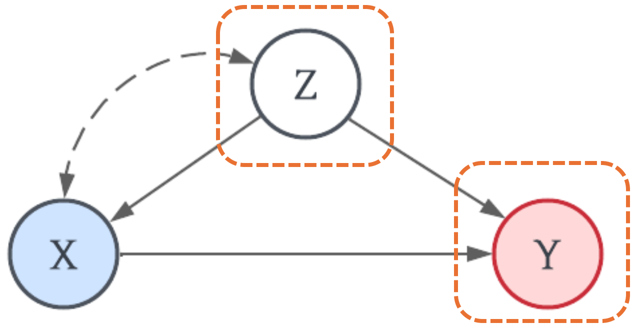
We propose ID-GEN
a sampling version of the Identification algorithm
for semi-Markovian causal models
using conditional generative models.

ID-GEN: Proposed Approach

- First, we decompose the interventional sampling problem into multiple sub-problems based on c-components.

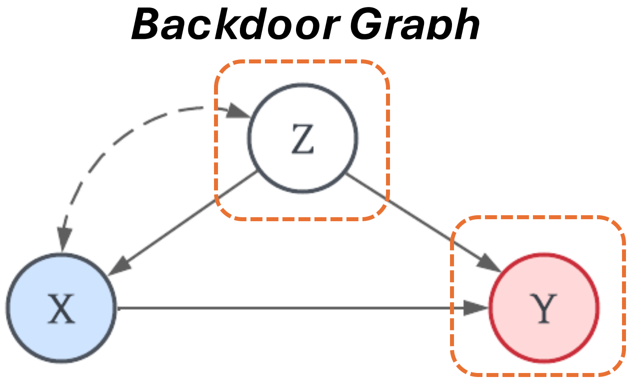
$$\textbf{if } C(G \setminus \mathbf{X}) = \{S_1, \ldots, S_k\} \textbf{ then } \{\text{Step 4}\}$$
$$\quad \textbf{for } \boxed{\text{each } S_i} \in C(G \setminus \mathbf{X}) = \{S_1, \ldots, S_k\} \textbf{ do}$$
$$\qquad \mathcal{H}_i \textbf{=ID-GEN}(S_i, \mathbf{X} = \mathbf{V} \setminus S_i, G, \hat{\mathbf{X}}, \hat{G}, \mathcal{D})$$

Backdoor Graph



$$P(y|do(x)) = \sum_z \boxed{P(z|do(x))}\,\boxed{P(y|do(x,z))}$$

ID-GEN: Proposed Approach



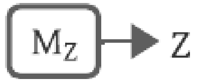
- Next, we train a set of conditional models for each factor.

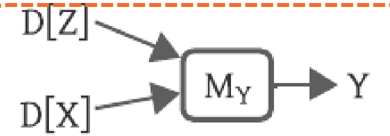
Algorithm 2 `ConditionalGMs`$(\mathbf{Y}, \mathbf{X}, G, \mathcal{D}, \hat{\mathbf{X}}, \hat{G})$

for each $V_i \in \{\mathbf{X} \cup \hat{\mathbf{X}}\}$ **do**
  Add node $(V_i, \emptyset)$ to $\mathcal{H}$ {Initialized $\mathcal{H} = \emptyset$}
for each $V_i \in \mathbf{Y}$ in the topological order $\pi_{\hat{G}}$ **do**
  Let M_{V_i} be a model trained on $\mathcal{D}[V_i, V_\pi^{(i-1)}]$ such that $M_{V_i}(V_\pi^{(i-1)}) \sim P(v_i | v_\pi^{(i-1)})$
  Add node (V_i, M_{V_i}) to $\mathcal{H}$
  Add edge $V_j \to V_i$ to $\mathcal{H}$ for all $V_j \in V_\pi^{(i-1)}$
Return $\mathcal{H}$.

$$P(z|do(x))$$

$$[Z] \sim P(Z)$$



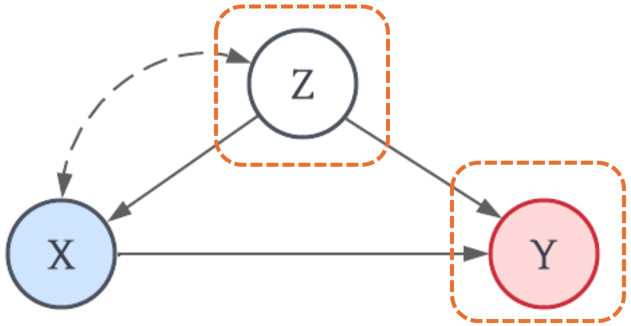


$$[Y] \sim P(Y|X, Z)$$

$$P(y|do(x, z))$$

ID-GEN: Proposed Approach

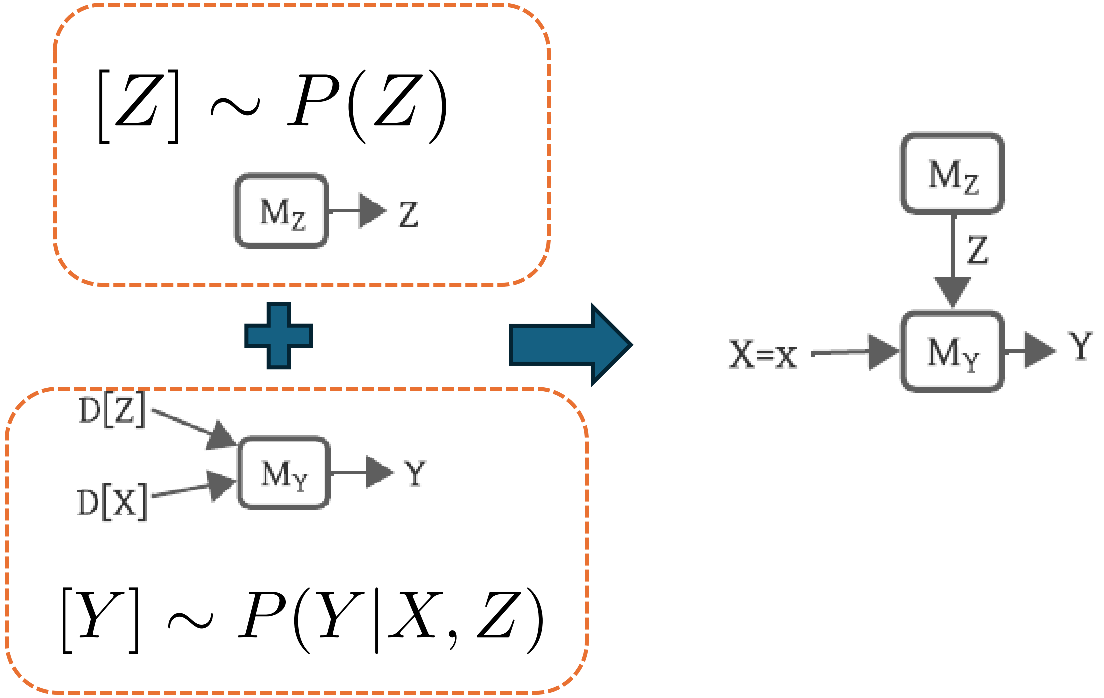
Backdoor Graph



- Finally, we connect them to build a neural network called sampling network.

Algorithm 3 $\texttt{MergeNetwork}(\{\mathcal{H}_i\}_{\forall i})$

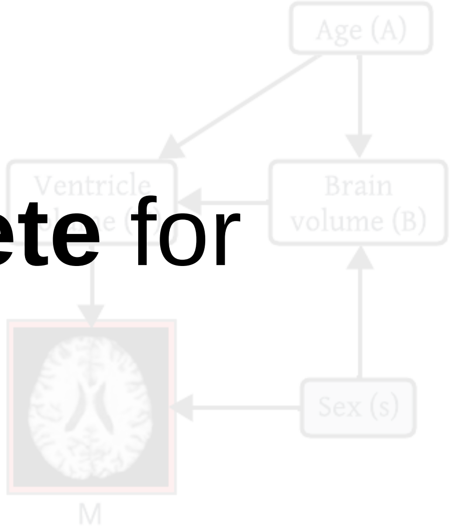
Input: Set of sampling networks $\{\mathcal{H}_i\}_{\forall i}$.
Output: A connected DAG sampling network $\mathcal{H}$.
for $H_i \in \{\mathcal{H}_i\}_{\forall i}$ **do**
  for $M_{V_j} \in \mathcal{H}_i$ **do**
    if $M_{V_j} = \emptyset$ and $\exists M_{V_k} \in \mathcal{H}_r, \forall r$ such that $V_j = V_k$ and $M_{V_k} \neq \emptyset$ **then**
      $M_{V_j} = M_{V_k}$
Return $\mathcal{H} = \{\mathcal{H}_i\}_{\forall i}$ {All $\mathcal{H}_i$ are connected.}



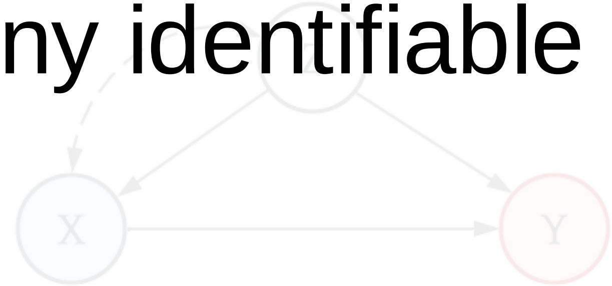
We can generate interventional samples!

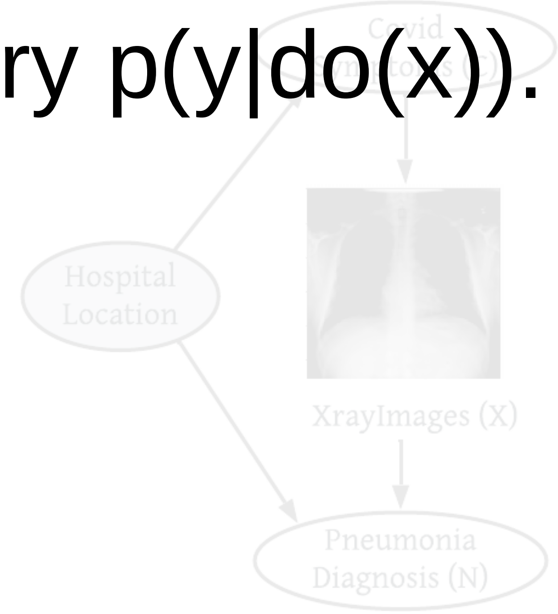
$$[Z, Y] \sim P(Z) * P(Y|X, Z)$$

Can we always do this?



Theorem: ID-GEN is **sound and complete** for any identifiable query p(y|do(x)).





Fairness: CelebA Image to Image Translation.

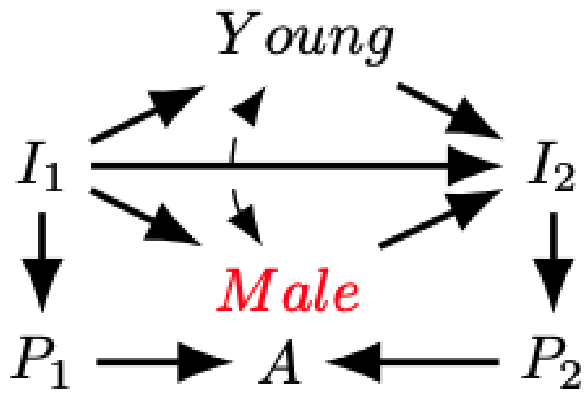
- Assess large generative models for the Male to the Female domain translation task.
- Translation: Causal or spurious?
- Correlation among different attributes learned by models.

CelebA Image to Image Translation:

- Original image I_1
- Edited image I_2 based on sex and age.
- All attributes of I_1 and I_2.
- A: new additional attributes (ex: Makeup)

- What is the causal effect of changing the Male domain to the Female domain on the appearance of a new attribute?

  $P(A|do(Male = 0))$.



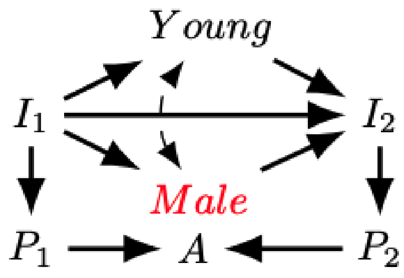
Conditional Generative Models are Sufficient to Sample from Any Causal Effect Estimand

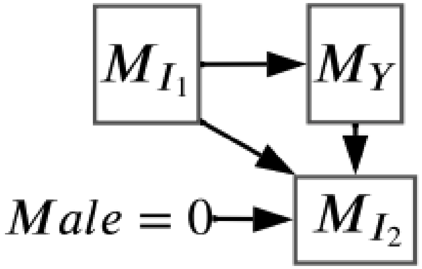
$$P(I_2|Do(Male = 0)) =$$

$$\int_{Young, I_1} P(I_2|Male = 0, Young, I_1)P(Young, I_1)$$

$$\boxed{I_1 \sim M_{I_1}}, \boxed{Young \sim M_Y(I_1)}$$

$$I_2 \sim M_{I_2}(Male = 0, Young, I_2)$$

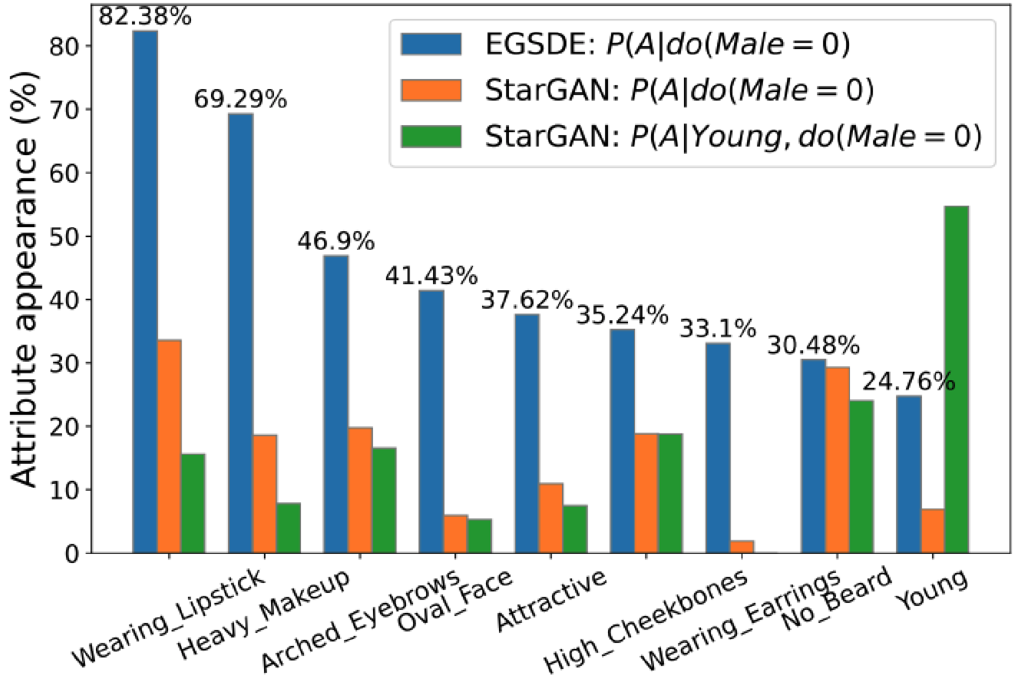




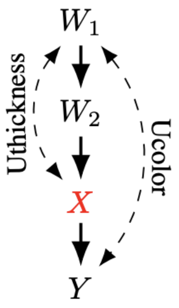
- For M_{I_2} use the following generative models:
  - EGSDE [4]
  - StarGAN [5]

Observations

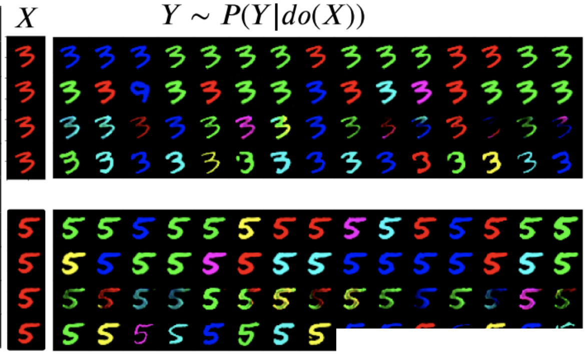
- EGSDE adds
  - Causal
    - WearingLipstick attribute to 82%.
    - HeavyMakeup: 69.28%
  - Non-causal
    - Attractive(37.61%) ?
    - Young(24.76%) ?



ID-GEN for Spurious Correlation & Explainability

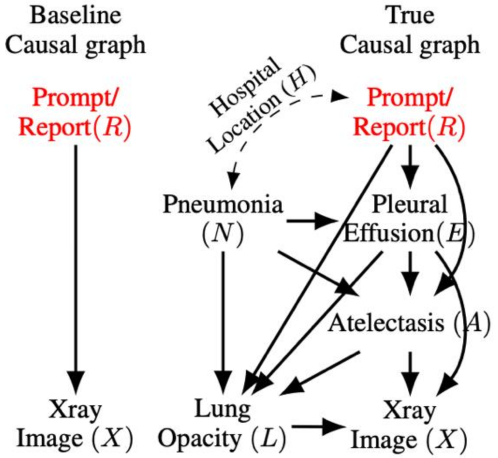


		FID ↓
$X=3$	Cond	50.83
	DCM	66.57
	NCM	61.11
	Ours	**25.66**
$X=5$	Cond	41.35
	DCM	60.61
	NCM	71.50
	Ours	**22.67**



Color	True $P_x(y)$	Cond	DCM	**Ours**
R	0.17	0.17	0.24	**0.13**
G	0.17	0.45	0.27	**0.24**
B	0.17	0.15	0.28	**0.21**
Yw	0.17	0.19	0.05	**0.14**
Mg	0.17	0.02	0.07	**0.11**
Cy	0.17	0.10	0.09	**0.18**
TVD	0	0.54	0.58	**0.25**

Figure 4: (Left:) Causal graph with color and thicl
the better) of each algorithm and images generatec
the $P_x(y)$ images generated by each algorithm. W



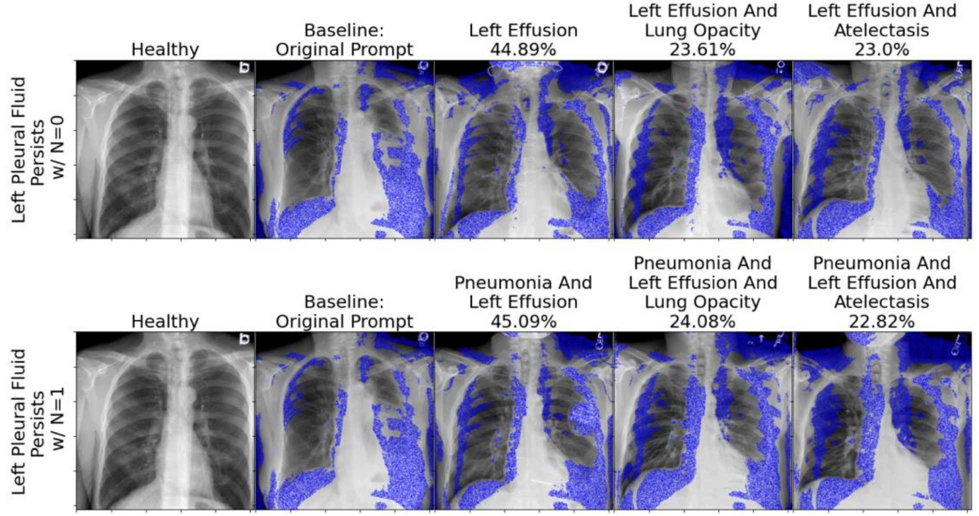


Figure 6: Left: Baseline vs our causal graph. Right: images for specific prompt w/ and w/o pneumonia. Inferred attributes are shown with their likelihood. Blue indicates changes compared to healthy.

Takeaway!

- Given observational data and a causal graph,
- Conditional generative models are **indeed** sufficient to sample from any causal effect estimand.

- Codes are available at: github.com/musfiqshohan/IDGEN

Thank you!

References

- [1] Shpitser, Ilya, and Judea Pearl. "Complete identification methods for the causal hierarchy." *Journal of Machine Learning Research* 9 (2008): 1941-1979
- [2] Ho, Jonathan, and Tim Salimans. "Classifier-free diffusion guidance." *arXiv preprint arXiv:2207.12598* (2022).
- [3] Jung, Yonghan, Jin Tian, and Elias Bareinboim. "Learning causal effects via weighted empirical risk minimization." *Advances in neural information processing systems* 33 (2020): 12697-12709.
- [4] Ribeiro, Fabio De Sousa, et al. "High Fidelity Image Counterfactuals with Probabilistic Causal Models." (2023).
- [5] Pearl, Judea. "The do-calculus revisited." arXiv preprint arXiv:1210.4852 (2012).
- [6] Report to Xray generation: Chambon, Pierre, et al. "Roentgen: Vision-language foundation model for chest x-ray generation." arXiv preprint arXiv:2211.12737 (2022).
- [7] Report to symptoms extraction: Gu, Jawook, et al. "CheX-GPT: Harnessing Large Language Models for Enhanced Chest X-ray Report Labeling." arXiv preprint arXiv:2401.11505 (2024).
- [8] Invariant Prediction: Subbaswamy, Adarsh, Peter Schulam, and Suchi Saria. "Preventing failures due to dataset shift: Learning predictive models that transport." The 22nd International Conference on Artificial Intelligence and Statistics. PMLR, 2019.
- [9] Lee, Kenneth, Md Musfiqur Rahman, and Murat Kocaoglu. "Finding invariant predictors efficiently via causal structure." Uncertainty in Artificial Intelligence. PMLR, 2023.