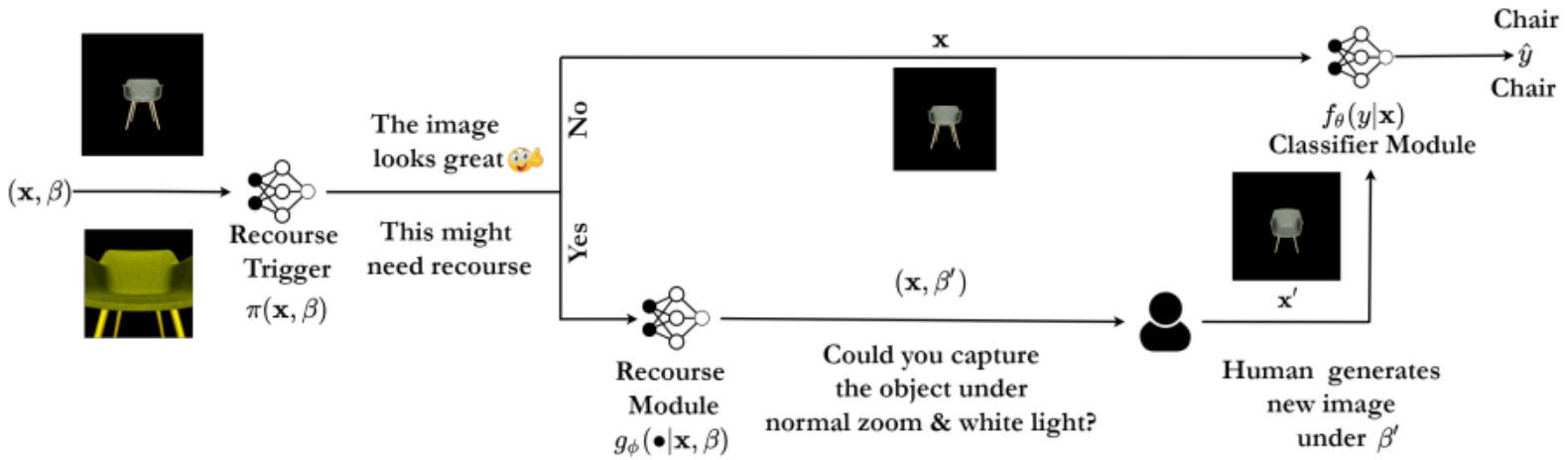
Learning Recourse on Instance Environment to Enhance Prediction Accuracy

- Lokesh N, Sai Koushik, Abir De, Sunita Sarawagi



Classifier:
Do not train on the subset that will be recoursed during test time.

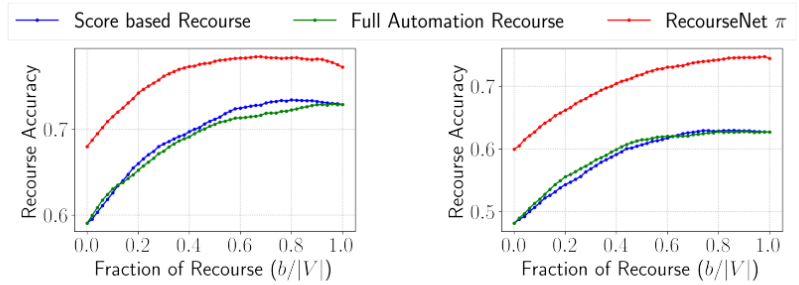
$$\max_{\theta,\phi,\pi} \sum_{\substack{i\in D \\ j\in B}} \log \Big[(1-\pi(\mathbf{x}_{ij}, \boldsymbol{\beta}_{ij}))\, f_\theta(y_i \mid \mathbf{x}_{ij}) + \pi(\mathbf{x}_{ij}, \boldsymbol{\beta}_{ij}) f_\theta(y_i \mid Z(z_i, \operatorname{argmax}_{\boldsymbol{\beta}} g_\phi(\boldsymbol{\beta} \mid \mathbf{x}_{ij}, \boldsymbol{\beta}_{ij}))) \Big]$$

subject to, $\sum_{i\in D, j\in B} \pi(\mathbf{x}_{ij}) \le b$,

$\pi(\mathbf{x}_{ij}, \boldsymbol{\beta}_{ij}) \in \{0, 1\}$

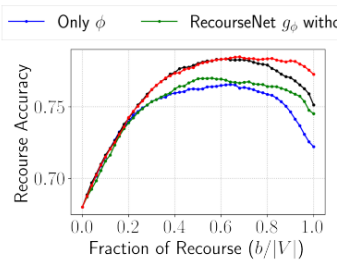
Recourse Module:
Hallucinate the counterfactual accuracy achievable on the recourse instances and use it while training.

Recourse Trigger:
Recourse only if the accuracy improves on the recoursed instance by atleast a δ

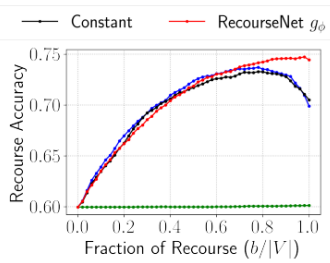


(a) Shapenet-Large

(b) Shapenet-Small



(a) Shapenet-Large.



(b) Shapenet-Small.