Predicting Drug Sensitivity of Cancer Cell Lines via Collaborative Filtering with Contextual Attention

Yifeng Tao[1,2,#], Shuangxia Ren[3,4,#], Michael Q. Ding[3], Russell Schwartz[1,5,*], Xinghua Lu[3,4,6,*]

[1]Computational Biology Department, School of Computer Science, Carnegie Mellon University

[2]Joint Carnegie Mellon-University of Pittsburgh Ph.D. Program in Computational Biology
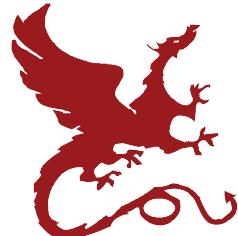
[3]Department of Biomedical Informatics, School of Medicine, University of Pittsburgh

[4]Intelligent Systems Program, School of Computing and Information, University of Pittsburgh

[5]Department of Biological Sciences, Carnegie Mellon University
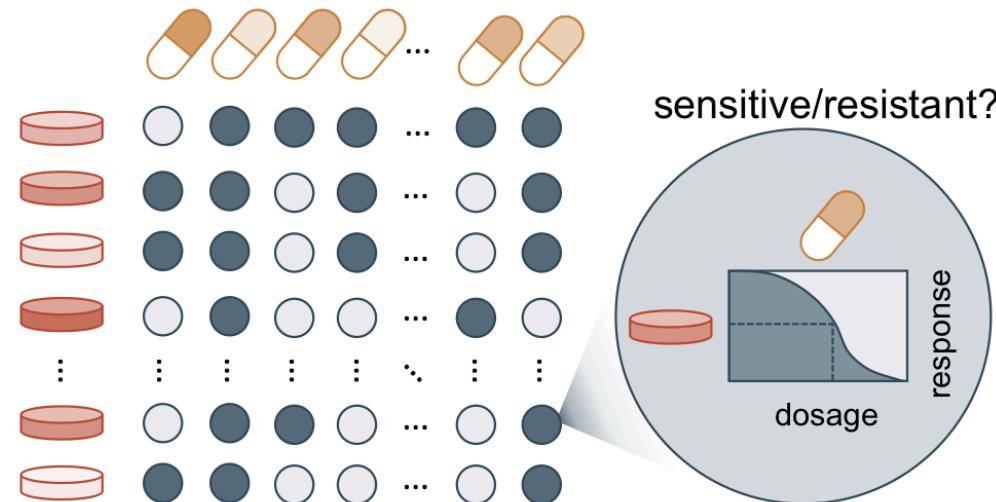
[6]Department of Pharmaceutical Science, School of Medicine, University of Pittsburgh

Anti-cancer drug recommendation

- Tumor resistance to drugs is complex
  - Both inter- and intra- tumor heterogeneities (Schwartz and Schäffer, 2017) .
  - Patients of same cancer type may have distinct prognoses (Priedigkeit et al. 2017) .

- Large scale cancer cell line assays
  - NCI-60 (Shoemaker 2006), CCLE (Barretinna et al. 2012), GDSC (Yang et al. 2013) etc.
  - Screening of cell line resistance to a panel of potential drugs.
 - Molecular profiles of cell lines.
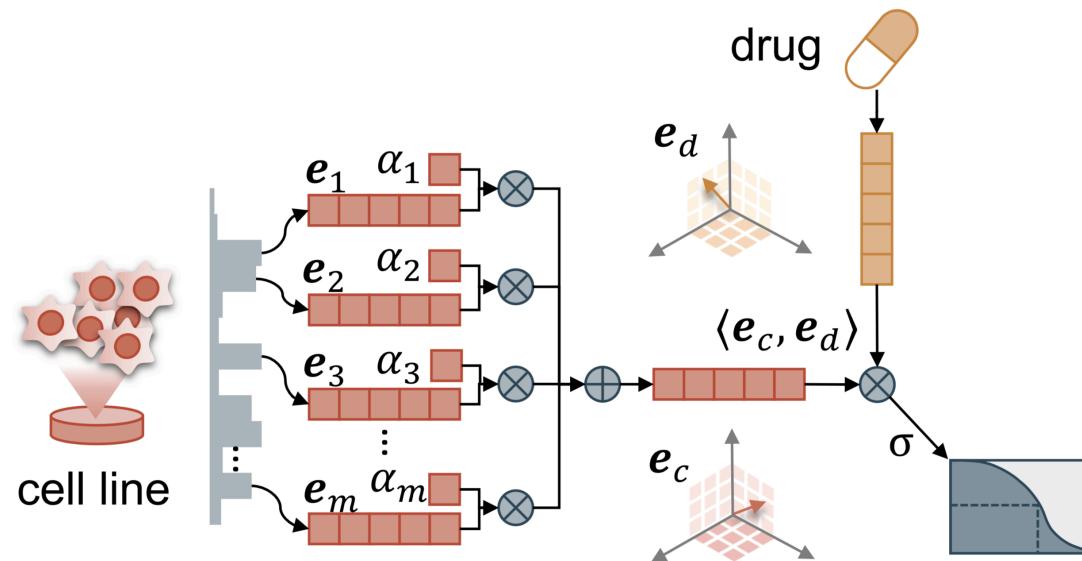
Existing work and challenges
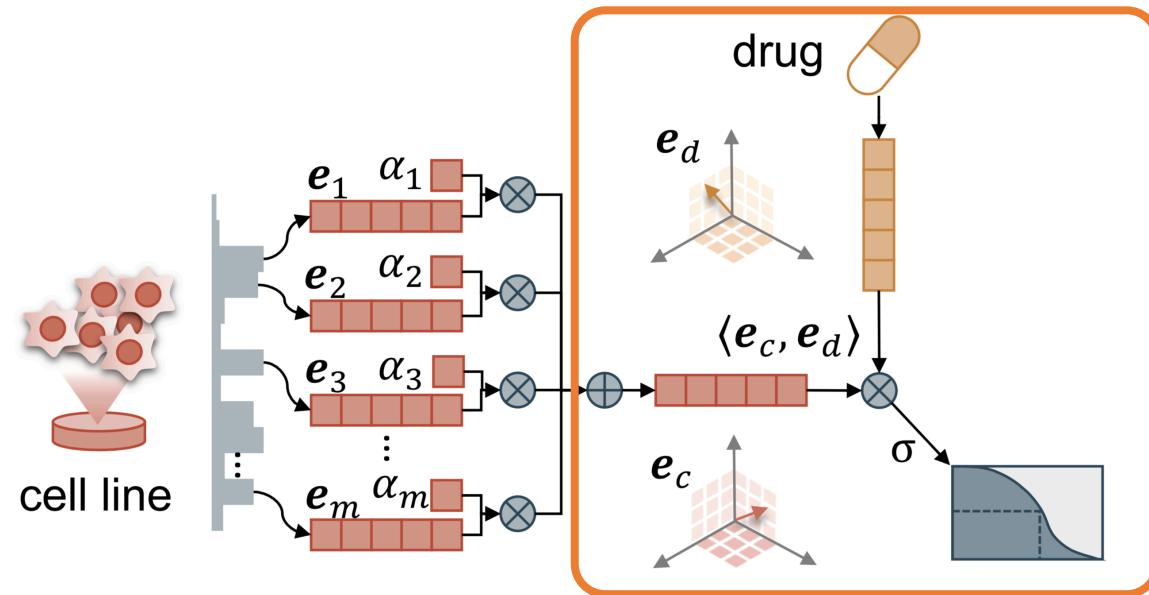
- Existing methods
  - Classical machine learning models: Elastic net (Yuan et al. 2016), Bayesian (Gonen and Margolin 2014) etc.
  - Cell line or drug similarity: Network (Wei et al. 2019), collaborative filtering (Liu et al. 2018) etc.
  - Deep learning models: MLP (Ding et al. 2018), DeepDR (Chiu et al. 2019) etc.
- Challenges in predicting drug response of cancer cell lines
  - Robustness: noise.
  - Contextual effects: gene interactions.
  - Interpretability: biomarkers.

Methods

- CADRE: Contextual Attention-based Drug REsponse
  - Collaborative filtering: copes with noisy data.
  - Attention mechanism: improves interpretability and performance.
  - Pretrained gene embeddings: boosts performance further.

Overall architecture: Collaborative filtering

$$\hat{y}_{c,d} = \sigma\left(\langle \boldsymbol{e}_c, \boldsymbol{e}_d \rangle\right) = \frac{1}{1 + \exp(-\boldsymbol{e}_c^\mathsf{T} \boldsymbol{e}_d)}$$

$$\ell(\hat{y}_{c,d}, y_{c,d}; \mathcal{W}) = \mathrm{CrossEnt}(\hat{y}_{c,d}, y_{c,d}) + \frac{\lambda_2}{2} \cdot \ell_2(\mathcal{W})$$

$$\boldsymbol{e}_c = \sum_{i=1}^{m} 1 \cdot \boldsymbol{e}_i = 1 \cdot \boldsymbol{e}_1 + 1 \cdot \boldsymbol{e}_2 + \ldots + 1 \cdot \boldsymbol{e}_m$$
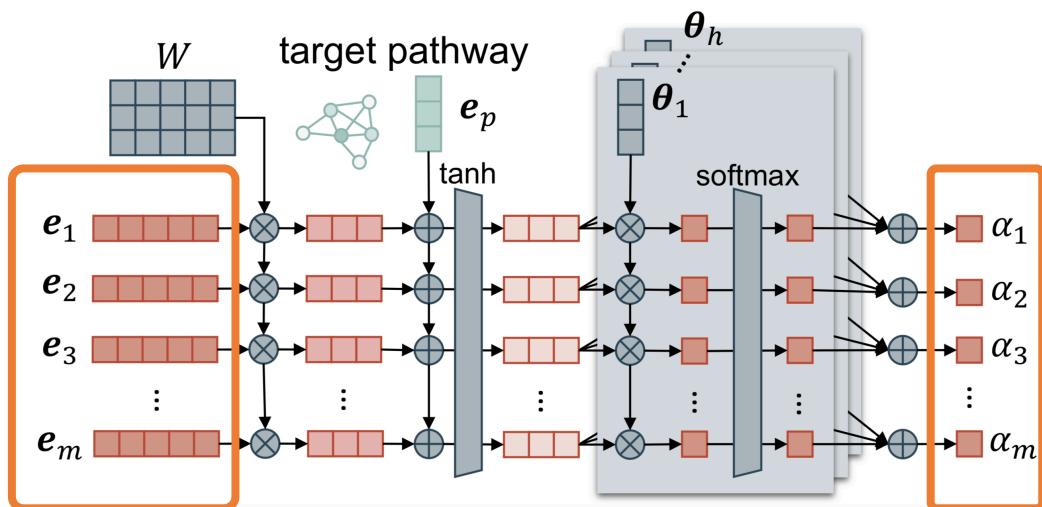
SADRE: Self-Attention-based Drug REsponse

- Cell embedding is the weighted sum of gene embeddings:

$$e_c = \sum_{i=1}^{m} \alpha_i \cdot e_i = \alpha_1 \cdot \mathbf{e}_1 + \alpha_2 \cdot \mathbf{e}_2 + ... + \alpha_m \cdot \mathbf{e}_m$$

$$\alpha_1, \alpha_2, ..., \alpha_m = \text{Self-Attention}(e_1, e_1, ..., e_m)$$

- Self-attention implemented as a sub-neural network:

$$\beta_{i,j} = \boldsymbol{\theta}_j^\mathsf{T} \tanh(W e_i), \qquad i = 1, 2, ..., m, \ j = 1, 2, ..., h$$
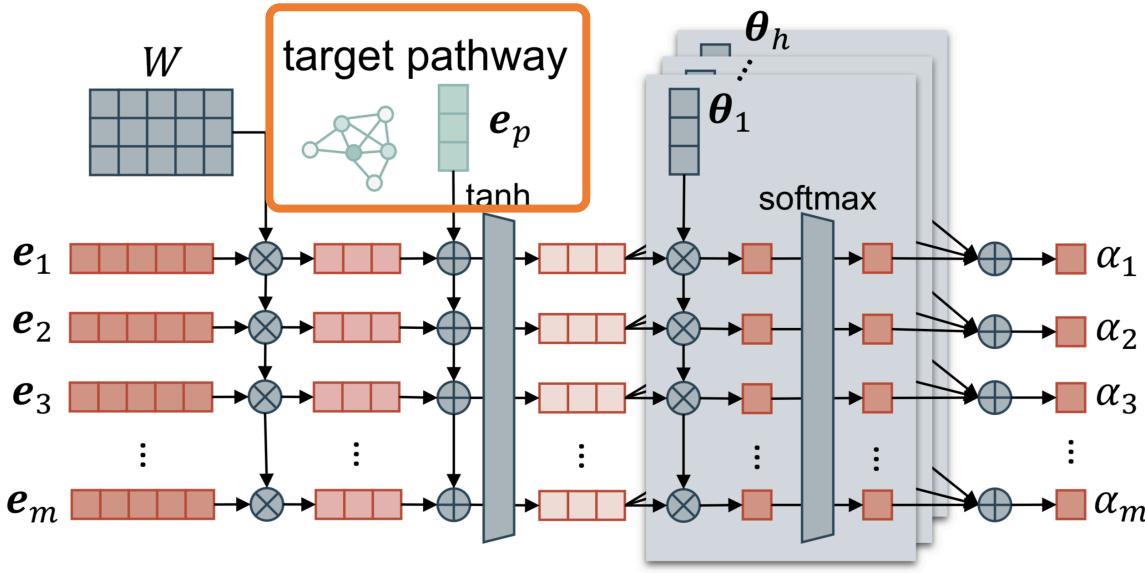
$$\alpha_{1,j}, \alpha_{2,j}, ...\alpha_{m,j} = \text{softmax}(\beta_{1,j}, \beta_{2,j}, ...\beta_{m,j}), \qquad j = 1, 2, ..., h$$

$$\alpha_{i,j} = \exp(\beta_{i,j}) \bigg/ \sum_{i'=1}^{m} \exp(\beta_{i',j}), \qquad i = 1, 2, ..., m$$

$$\alpha_i = \sum_{j=1}^{h} \alpha_{i,j} = \alpha_{i,1} + \alpha_{i,2} + ... + \alpha_{i,h}, \qquad i = 1, 2, ..., m$$

CADRE: Contextual Attention-based Drug REsponse

- Drug pathway knowledge is integrated.
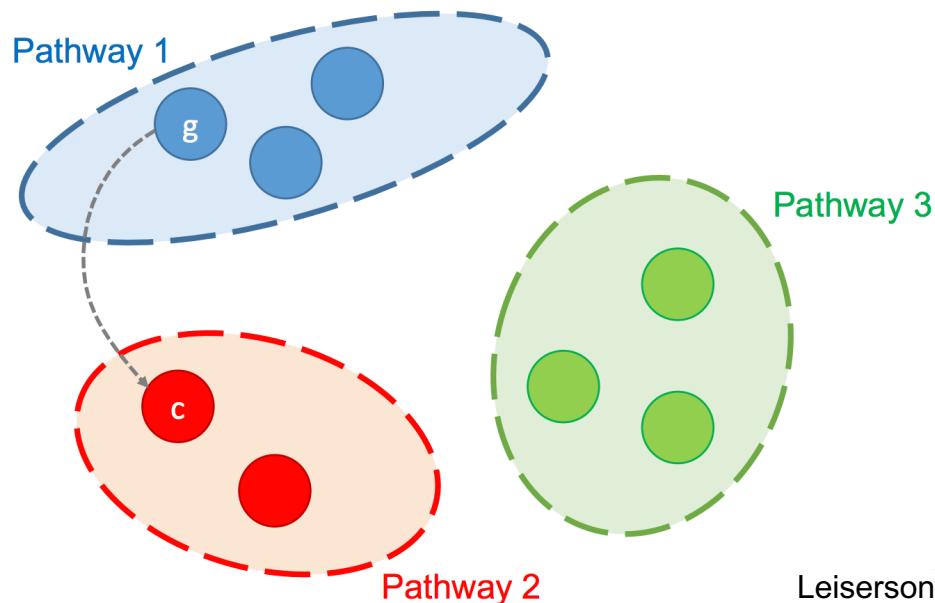
$$\alpha_1, \alpha_2, ..., \alpha_m = \text{Contextual-Attention}\left(\boldsymbol{e}_1, \boldsymbol{e}_1, ..., \boldsymbol{e}_m, \boldsymbol{e}_p\right)$$

$$\beta_{i,j} = \boldsymbol{\theta}_j^{\mathsf{T}} \tanh(W\boldsymbol{e}_i + \boxed{\boldsymbol{e}_p}), \qquad i = 1, 2, ..., m, \ \ j = 1, 2, ..., h$$
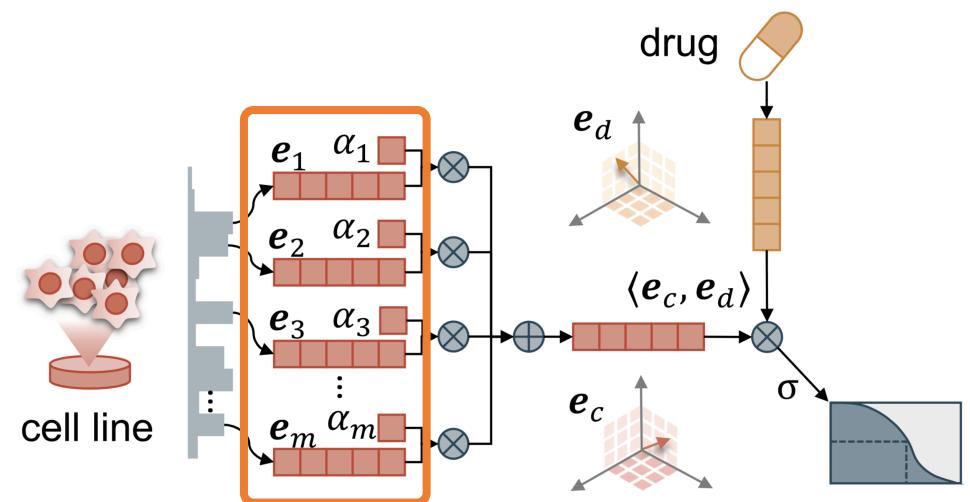
Pretraining gene embeddings

- Gene embedding pretrained using gene2vec, a variant of word2vec, on GEO.
- Co-occurrence pattern.

$$\Pr\left(c \in \text{Context}(g) \mid g\right) = \frac{\exp\left(\mathbf{e}_g^\mathsf{T} \mathbf{v}_c\right)}{\sum_{c' \in \mathcal{G}} \exp\left(\mathbf{e}_g^\mathsf{T} \mathbf{v}_{c'}\right)}$$

Pathway 1

Pathway 3

Pathway 2
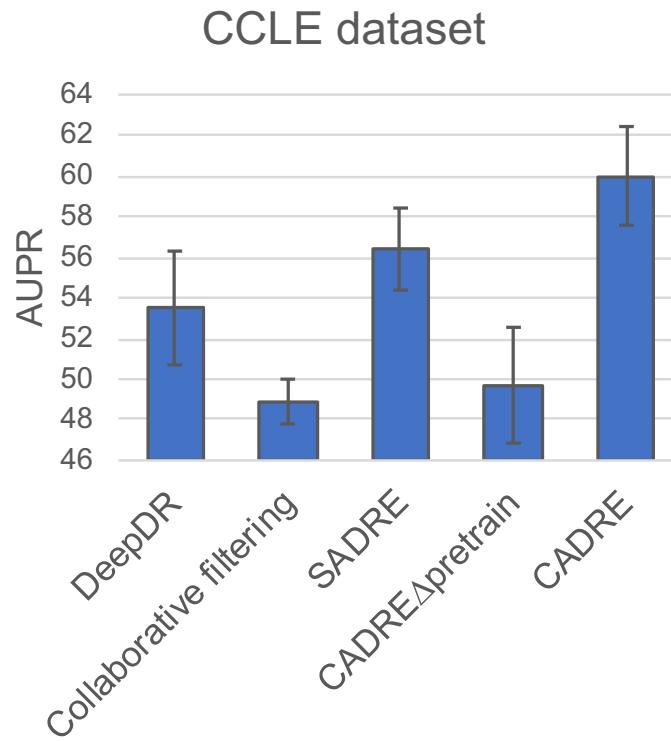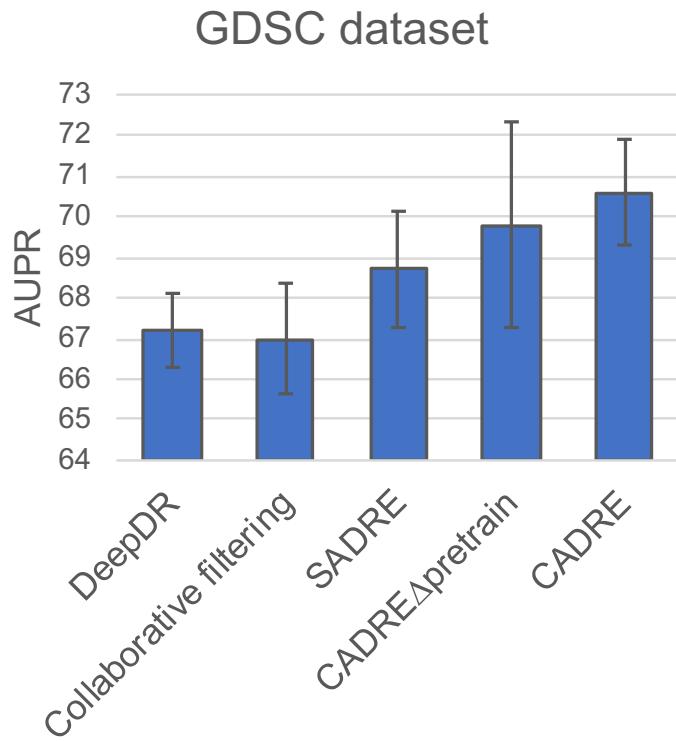
Leiserson et al. 2015;
Mikolov et al. 2013

drug

$\boldsymbol{e}_d$

cell line

$\boldsymbol{e}_1$ α_1

$\boldsymbol{e}_2$ α_2

$\boldsymbol{e}_3$ α_3

$\boldsymbol{e}_m$ α_m

$\langle \boldsymbol{e}_c, \boldsymbol{e}_d \rangle$

$\boldsymbol{e}_c$

σ
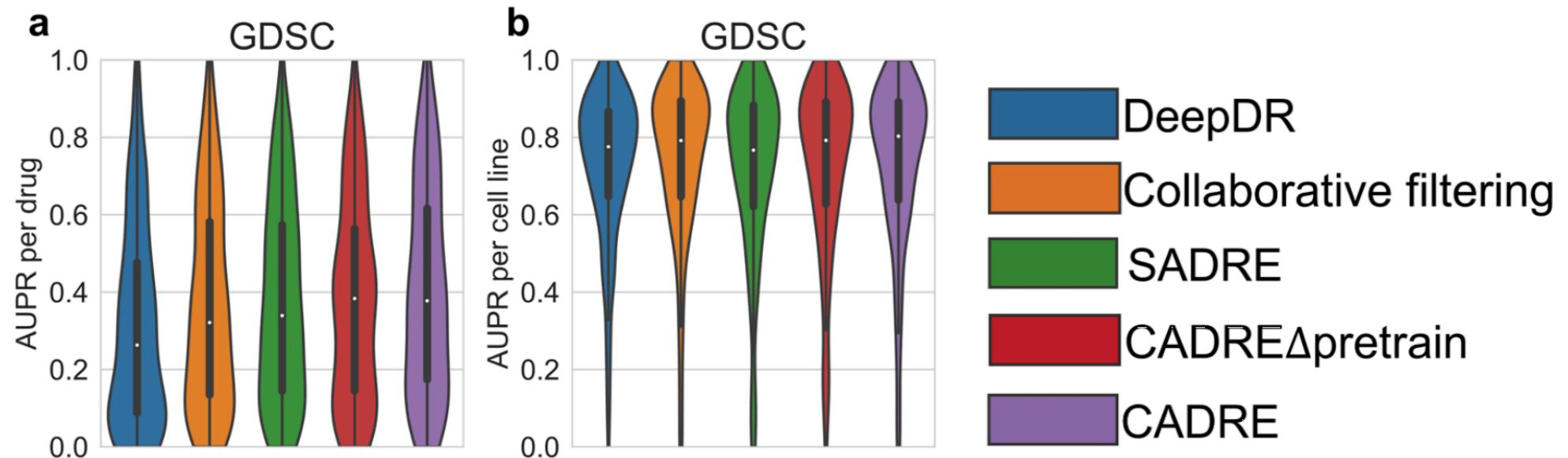
Results: Performance

- Outperforms competing models


GDSC dataset

CCLE dataset
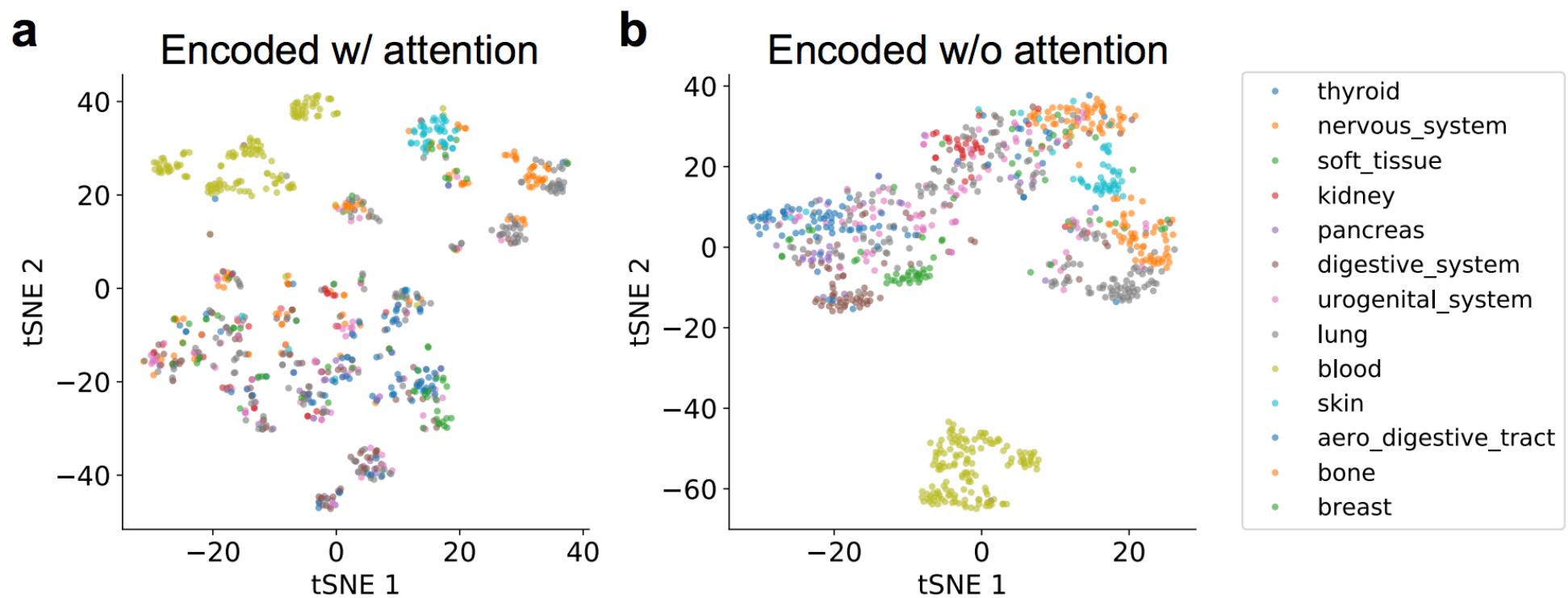
Effective attention-encoded cell line representation

- Major improvements from AUPR per drug…
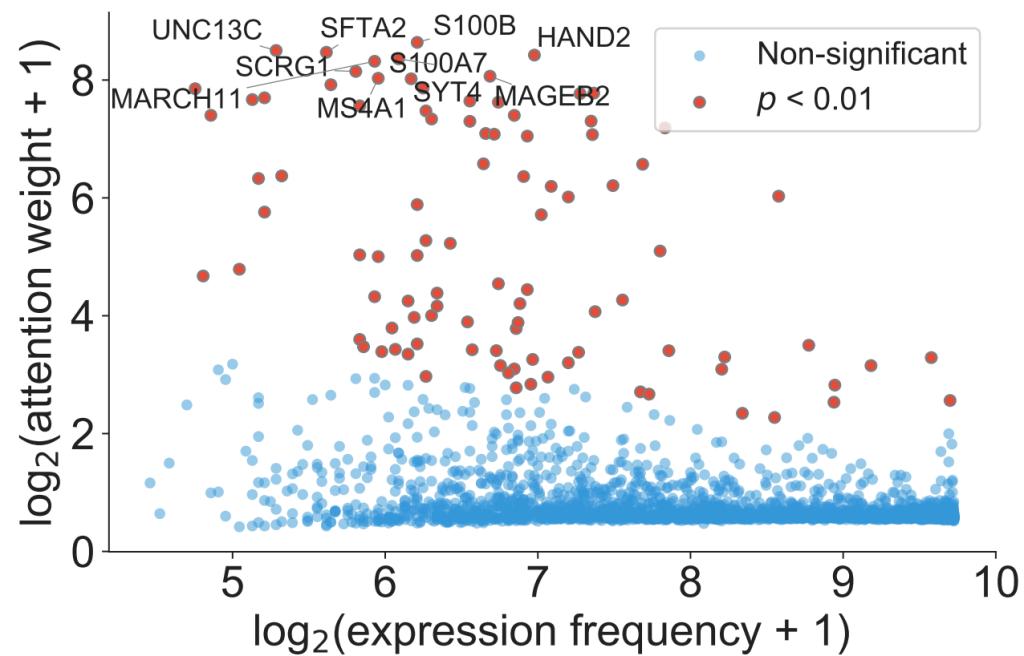
Effective attention-encoded cell line representation

Identifies critical biomarkers related to drug resistance

- CADRE identifies critical biomarkers related to drug resistance
- Two enriched pathways
  - Export from cell
  - Signaling receptor binding

Conclusions and future directions

- Conclusions:
    - CADRE integrates the attention mechanism into the collaborative filtering framework.
    - Outperforms competing models in predicting drug responses from RNA profiles of cell lines.
    - Effective attention-encoded cell line representation.
    - Identifies critical biomarkers related to drug resistance.

- Future directions
    - Drug recommendation *in vivo*: intra-tumor heterogeneity.
    - Better drug feature representation.

Acknowledgment

- Prof. Xinghua Lu
- Prof. Russell Schwartz
- Shuangxia Ren
- Michael Q. Ding