a. **What is the problem the article's trying to address?**

Predictive models are of great use in cancer research because they allow to estimate cancer therapy possible response given patient's genetic data. But for scRNA-seq, there is not a lot of training data available, which affects quality of ML models.

b. **What are the related works in the field and why there is still a need to propose new solutions?**

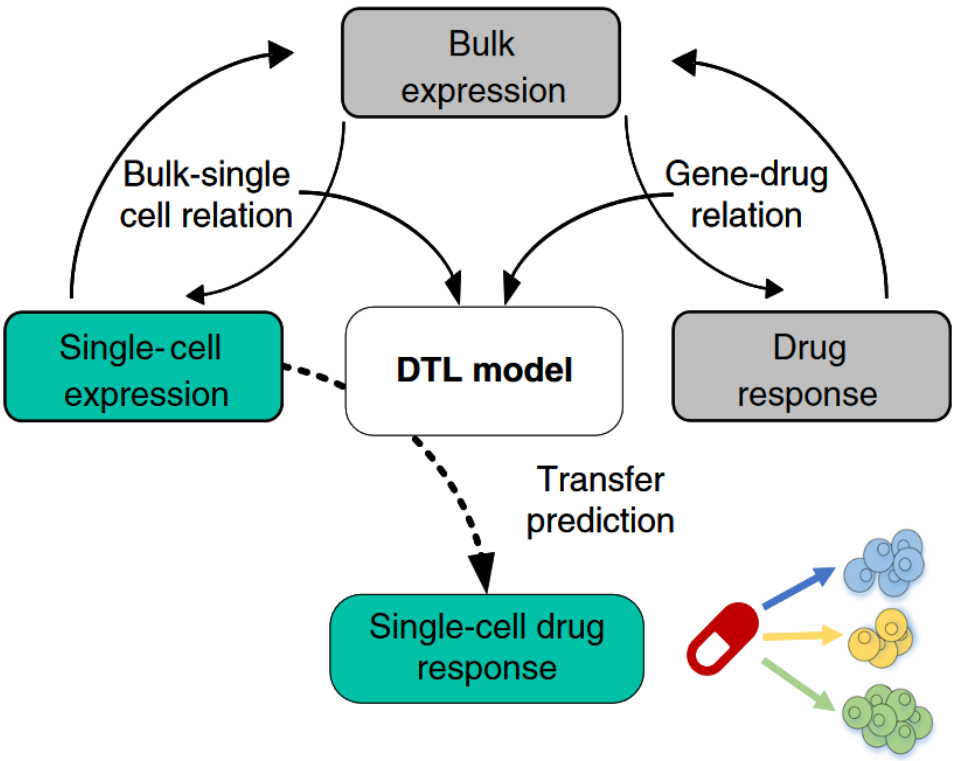
As authors mention, DL technologies are already widely applied to clustering, expression imputation, and batch effect correction. Deep Transfer Learning is already applied in combining multiple bulk RNA-seq datasets to predict drug response. However, there were no tool to directly use bulk RNA-seq data for predicting response of scRNA-seq input.

c. **What do the authors propose? Describe their solution, input data, processing, metrics etc.**

Authors apply Deep Transfer Learning (DTL) to build relations between bulk and singe cell levels, so that bulk data can be applied to predict single-cell response.

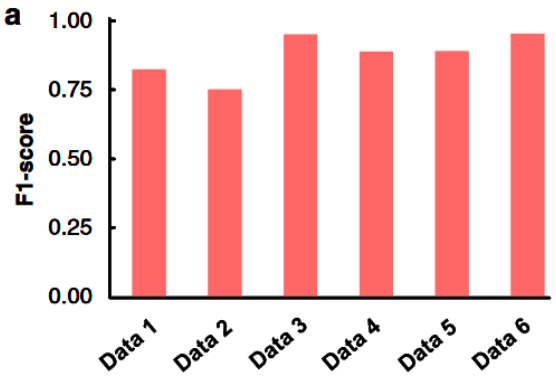
Not diving into extreme details, model (named scDEAL - single-cell Drug rEsponse AnaLysis) does following:

- learns gene-drug relation from bulk data of Genomics of Drug Sensitivity in Cancer (GDSC) and Cancer Cell Line Encyclopedia CCLE
- learns SC-gene relation using bulk and SC data
- connects gene-drug and gene-cell relation, which allows to predict cell-drug relation

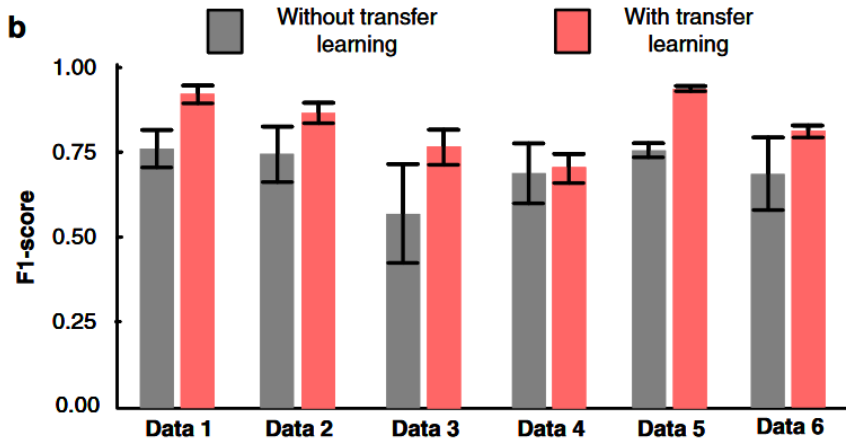


d. What are the major results and achievements of the proposed solution? How do they relate to the existing methods and what are the limitations?

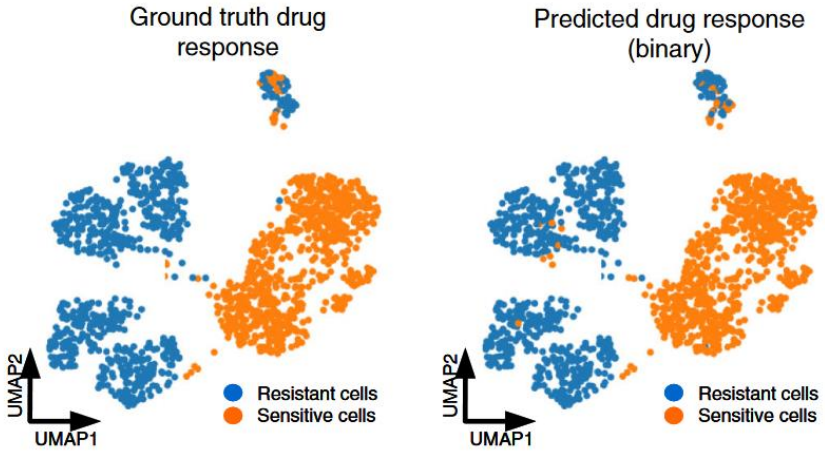
Resulting model performs quite accurate with mean F1 score of 0.89 in six datasets.



As a comparison with an existing method, authors tested their model with the models that just use same datasets without transfer learning, and in all cases transfer learning outperformed other models.



In an example, accuracy in predicting drug resistance in leukemia cells under I-BET treatment exceeds 95%. Authors also show that the model can predict specific genes responsible for drug response.



As for limitations, authors mention that prediction accuracy can vary because not all cell lines are well represented in training data collections, but they will update it. Another challenge is cross-species predictions, which are currently not optimized due to small amount of annotated animal single cell data.

e. What are the conclusions?

DTL approach is a very promising, as not only it provides a more accurate predictions for scRNA-seq data already, but also can be extended for another applications in fields where amount of training data is insufficient, not limiting with only drug-related data. Such fields are cell reprogramming, predicting CRISPR response, and many others.