Improvements for variational autoencoders inspired by biological abstractions

STUDENT: Qian-Wu Liao

MATRIKELNUMMER: 3631657

QianWu.Liao@stud.uni-heidelberg.de

SUPERVISOR: Daria Doncevic

SUPERVISOR: Dr. Carl Herrmann

Background

Deep learning-based methods such as autoencoders (AEs) [1] have been proven to be an effective technique to handle high-dimensional data. Unlike principal component analysis (PCA) that is limited by its linear nature, AEs are capable of capturing more complex patterns in data. Generally, AEs consist of an encoder that can learn a non-linear transformation compressing high-dimensional data into lower-dimensional vectors forming a latent space and a decoder that will then take the compressed vectors from the latent space and reconstruct it to the original dimensions. However, it is challenging to use AEs as generative models due to their discrete latent spaces. To this end, variational autoencoders (VAEs) [2] were developed. VAEs map high-dimensional data to a probability distribution where the lower-dimensional representations can be sampled, which better generalizes models. Even though VAEs can generate fairly complete and informative latent spaces and make accurate predictions, yet, these latent spaces provide little to no interpretability that is crucial for understanding biological mechanisms. To gain more interpretability of internal networks, prior biological information can be integrated into network structures. Ma et al. (2018) has successfully embedded the hierarchies of molecular subsystems about cellular processes in the network architecture, which models not only functional outcomes but also the mechanisms resulting in these outcomes.

Previous work

Based on the idea mentioned above, Seninge et al. (2021) proposed a novel VAE inspired by biological abstractions, VEGA (VAE Enhanced by Gene Annotations), whose architecture consists of a deep non-linear encoder and a single-layer masked linear decoder. An interpretable latent space and decoder

wiring of the VEGA is specified by *a priori* biological abstractions, e.g., gene regulatory networks. Latent variables and decoder connections can be initiated by master transcription factors (TFs) and TF-target gene connections and the weights of decoder connections offer direct interpretation of the relationships between the latent variables and the original features. Note that the decision of a single-layer decoder is to maximize the inference capacity of the encoder.

Aims of thesis

The VEGA efficiently provides interpretability of the latent space, which enables us to gain more biological insights. However, the single-layer masked decoder of the VEGA sacrifices the generative capacity for interpretability of the latent space. Therefore, we will mainly focus on improving this downside by inserting an additional layer (e.g., initiated by a list of second major TFs) in the decoder part to boost a predictive performance while maintaining biological interpretability.

25.10.21 - 07.11.21
Reproduce the results from the VEGA paper

Firstly, we will reproduce the results from the VEGA papaer to validate and understand the code provided by the author[1]. The VEGA is built using Python *PyTorch* framework [3] and the single-cell sequencing data is preprocessed by Python *SCANPY* toolkit [4].

08.11.21 - 05.12.21
Exploit the VEGA model to analyze the activity of the master regulators in the context of neuroblastoma transcriptomics data

Next, we will adapt the code to the neuroblastoma single-cell transcriptomics data [5] while using master TFs (MYCN and TFAP2B) and their gene regulatory networks to initiate the latent space and the decoder wiring.

06.12.21 - 30.01.22
Add an additional layer in the docoder part to improve a predictive performance

After adapting the VEGA model to the neuroblastoma data and evaluating the quality of the results, we will manage to add an additional layer (i.e., second major TFs) in the decoder part in an attempt to boost a predictive performance.

---

[1]`https://github.com/LucasESBS/vega`

31.01.22 - 27.02.22
Improve hard-coding decoder by regularization methods on weights

Hard-coding decoder connections using a binary matrix that describes the relationship between TFs and original genes impedes the model to correct prior information or explore unknow knowledge. We will take advantage of regularization methods on weights in the decoder part to make the model more robust.

28.02.22 - 03.04.22
Extend the usage of the VEGA model to bulk RNA-seq data

Last but not least, since the workflow of the VEGA is currently designed to cope with single-cell RNA sequencing (scRNA-seq) data only, we are expecting to extend the model to bulk RNA-seq data.

04.03.22 - 01.05.22
Buffer time: Summarize the work and write the thesis

References

[1] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313:504–507, 2006.

[2] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. *International Conference on Learning Representations (ICLR)*, 2014.

[3] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *NIPS-W*, 2017.

[4] F. Alexander Wolf, Philipp Angerer, and Fabian J. Theis. Scanpy: large-scale single-cell gene expression data analysis. *Genome Biology*, 19:15, 2018.

[5] Selina Jansky, Ashwini Kumar Sharma, Verena Körber, Andrés Quintero, Umut H. Toprak, Elisa M. Wecht, Moritz Gartlgruber, Alessandro Greco, Elad Chomsky, Thomas G. P. Grünewald, Kai-Oliver Henrich, Amos Tanay, Carl Herrmann, Thomas Höfer, and Frank Westermann. Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. *Nature Genetics*, 53:683–693, 2021.

[6] Jianzhu Ma, Michael Ku Yu, Samson Fong, Keiichiro Ono, Eric Sage, Barry Demchak, Roded Sharan, and Trey Ideker. Using deep learning to model the hierarchical structure and function of a cell. *Nature Methods*, 15:290–298, 2018.

[7] Lucas Seninge, Ioannis Anastopoulos, Hongxu Ding, and Joshua Stuart. Vega is an interpretable generative model for inferring biological network activity in single-cell transcriptomics. *Nature Communications*, 12:5684, 2021.