Contrastive Learning of Image and Structure-Based Representations in Drug Discovery

AIDD seminar - 08.06.2022

Ana Sanchez-Fernandez

Supervisor: Günter Klambauer

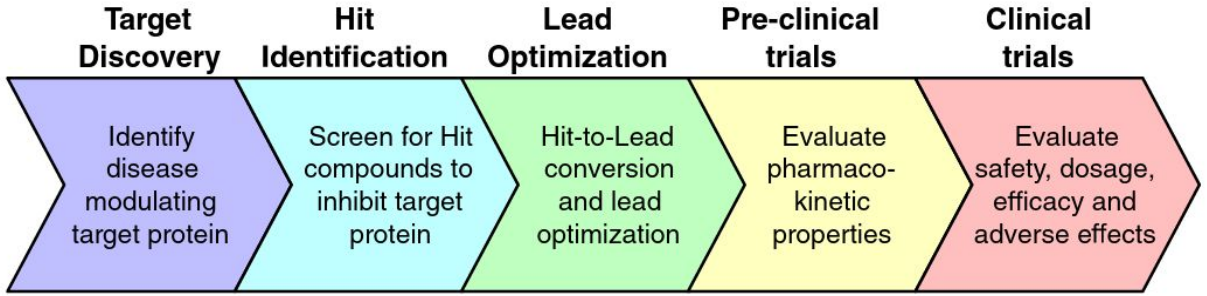
JOHANNES KEPLER UNIVERSITÄT LINZ

Overview

- Introduction
  - The drug discovery process
  - Fluorescence microscopy
  - Related work. ML for microscopy imaging
  - Self-supervised contrastive learning
- Method
- Results
- Conclusions
- Next steps

Introduction

The drug discovery process



- **Target-based:** starts by selecting a **protein** involved in certain **disease**
- **Molecules** are **screened** against that protein in order to find the ones with the desired **activity** (HTS)

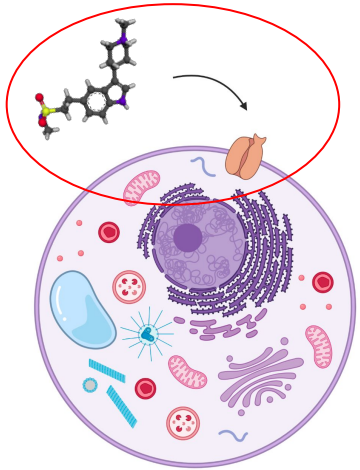
Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). **Principles of early drug discovery.** *British Journal Of Pharmacology, 162(6)*, 1239-1249. doi: 10.1111/j.1476-5381.2010.01127.x
Schaduangrat, N., Lampa, S., Simeon, S., Gleeson, M. P., Spjuth, O., & Nantasenamat, C. (2020). **Towards reproducible computational drug discovery.** *Journal of Cheminformatics 2020 12:1*, *12*(1), 1–30. https://doi.org/10.1186/S13321-020-0408-X

QSAR models

- HTS data has been successfully used in ML for drug discovery using **molecular structures** as input
- Limitations:
  - Chemical space is huge ⇨ Unfeasible to test every molecule
  - Biased data towards already known actives
  - Poor performance in insufficiently explored areas of the chemical space

Cumming, J. G., Davis, A. M., Muresan, S., Haeberlein, M., & Chen, H. (2013). **Chemical predictive modelling to improve compound quality.** *Nature Reviews Drug Discovery*, *12*(12), 948–962. https://doi.org/10.1038/NRD4128

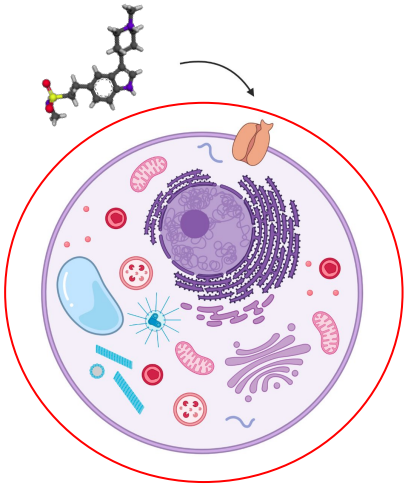
Fluorescence microscopy



- Microscopy images capture **morphological changes** produced by chemical compounds
- Provides information about molecules' **Mechanism of Action (MoA)**
- Allows clustering compounds by biological activity similarity

Niranj Chandrasekaran, S., Ceulemans, H., Boyd, J. D., & Carpenter, A. E. (2021). **Image-based profiling for drug discovery: due for a machine-learning upgrade?** *Nature Reviews Drug Discovery*. https://doi.org/10.1038/s41573-020-00117-w

Fluorescence microscopy

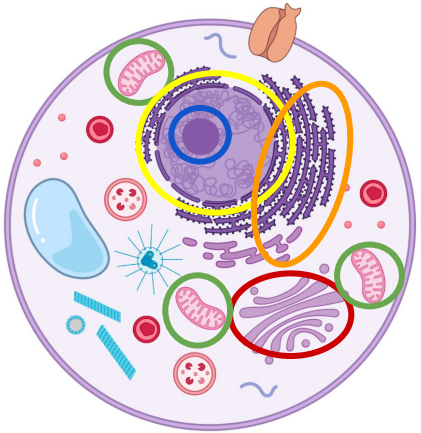


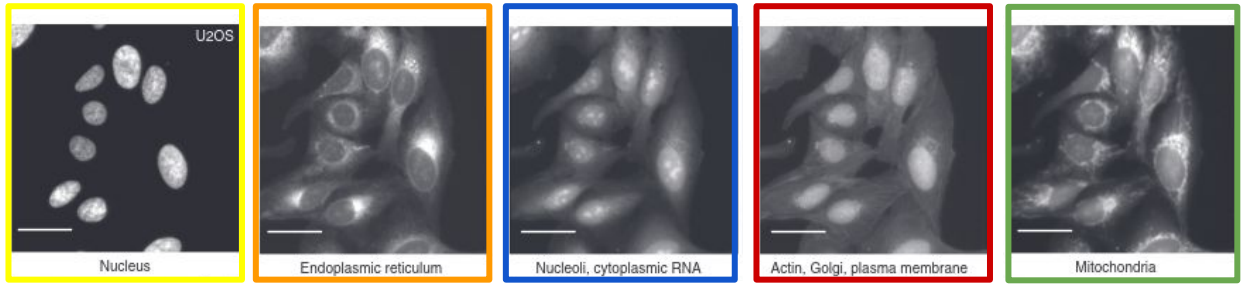
- Microscopy images capture **morphological changes** produced by chemical compounds
- Provides information about molecules' **Mechanism of Action (MoA)**
- Allows clustering compounds by biological activity similarity

Niranj Chandrasekaran, S., Ceulemans, H., Boyd, J. D., & Carpenter, A. E. (2021). **Image-based profiling for drug discovery: due for a machine-learning upgrade?** *Nature Reviews Drug Discovery*. https://doi.org/10.1038/s41573-020-00117-w

Cell Painting dataset

- **6 dyes** stain different **cell components**
- Displayed in **5 channels**
- Image resolution: 520x696

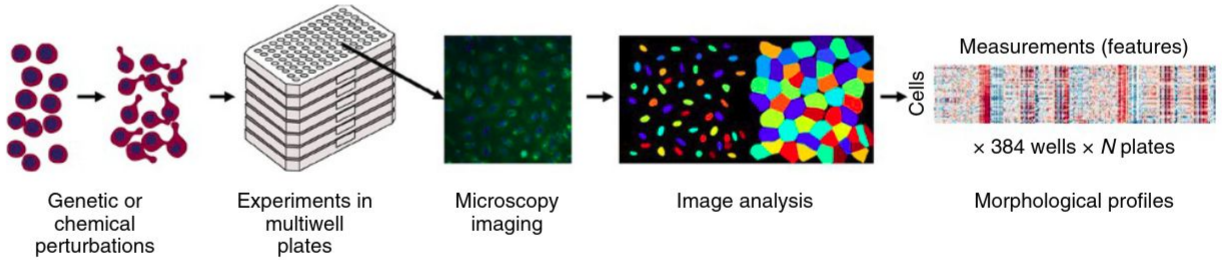




Bray, M. A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., …, & Carpenter, A. E. (2016). **Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes.** *Nature Protocols 2016 11:9*, *11*(9), 1757–1774. https://doi.org/10.1038/nprot.2016.105

Cell Painting dataset

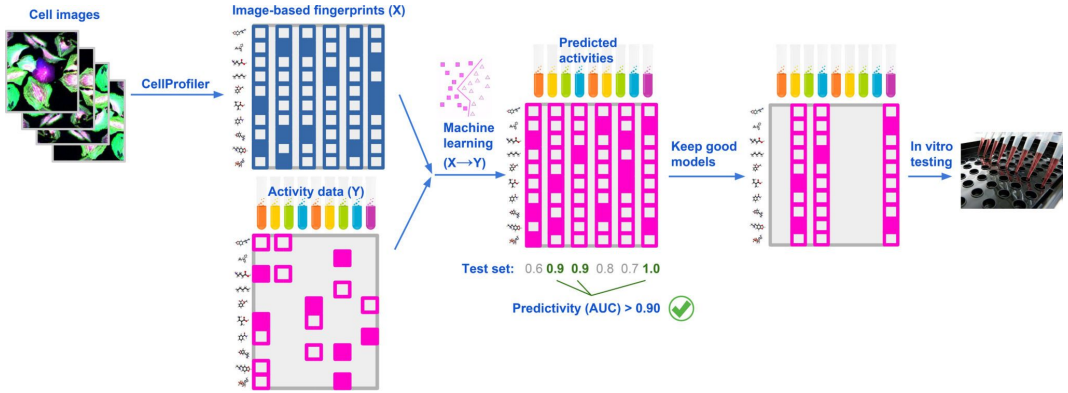
- 919,265 five-channel **microscopy images**
- Perturbed with 30,610 **chemical compounds**
- Includes **image-based fingerprint** extracted with **CellProfiler**



Bray, M. A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., …, & Carpenter, A. E. (2016). **Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes.** *Nature Protocols 2016 11:9*, *11*(9), 1757–1774. https://doi.org/10.1038/nprot.2016.105

Related work. Image profile-based

- **Features** extracted by CellProfiler as input

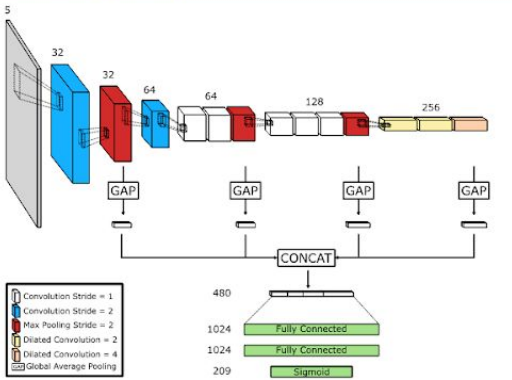


- Predicted active molecules tested in-vitro:
  - **280 and 50-fold enrichment** over initial hit rate

Simm, J., Klambauer, G., Arany, A., Steijaert, M., Wegner, J. K., Gustin, E., Chupakhin, V., Chong, Y. T., Vialard, J., Buijnsters, P., Velterm Ingrid, Vapirev, A., Singh, S., Carpenter, A. E., Wuyts, R., Hochreiter, S., Moreau, Y., & Ceulemans, H. (2018). **Repurposing high-throughput image assays enables biological activity prediction for drug discovery.** *Cell Chem Biol*, *25*(5), 611–618. https://doi.org/10.1016/j.chembiol.2018.01.015

Related work. CNNs for microscopy

- **Benchmark** different **CNN architectures** against MLP and image-based **fingerprint baseline**
- From **ChEMBL**, retrieve available **bioactivity** data for Cell Painting compounds
- **Multi-task** approach: 209 tasks (assays)
- CNNs outperformed baseline. Best model: high predictivity in 32% assays



Hofmarcher, M., Rumetshofer, E., Clevert, D., Hochreiter, S., & Klambauer, G. (2019). **Accurate Prediction of Biological Assays with High-Throughput Microscopy Images and Convolutional Networks.** *Journal Of Chemical Information And Modeling*, *59*(3), 1163-1171. doi: 10.1021/acs.jcim.8b00670

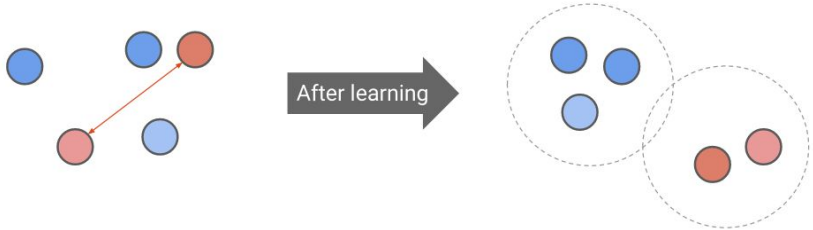
Self-supervised contrastive learning

- Fully **supervised** approaches have achieved outstanding results, but:
  - Generalization difficulties
  - Need large amounts of expensive labeled data
- **Self-supervised** learning: supervised learning objective for **unlabeled** data
- Ability to produce **transferable representations** to different tasks

Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., Tang, J., & Fellow, I. (2021) **Self-supervised Learning: Generative or Contrastive.** *ArXiv, 1911.05722.*

Self-supervised contrastive learning

- **Contrastive learning:** predict **relationship** between multiple samples
- Learn an embedding space in which similar (“positive”) sample pairs are close to each other and dissimilar (“negative”) ones are far apart



- **InfoNCE objective.**

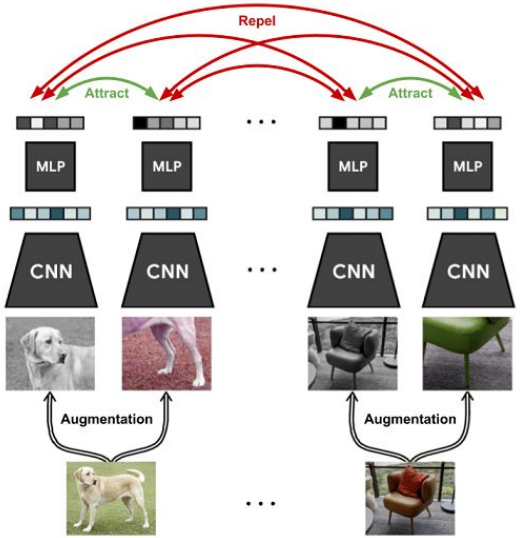
$$\mathrm{L}_{\mathrm{InfoNCE}} = -\ln \frac{\exp(\tau^{-1}\boldsymbol{x}_1^T\boldsymbol{y})}{\exp(\tau^{-1}\boldsymbol{x}_1^T\boldsymbol{y}) + \sum_{j=2}^{N}\exp(\tau^{-1}\boldsymbol{x}_j^T\boldsymbol{y})}$$

Weng, L. and Jong Wook Kim (2021,..). **Self-supervised learning. Self-prediction and contrastive learning.** [Tutorial presentation]. In *NeurIPS*, 2021.
Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., Tang, J., and Fellow, I. (2021) **Self-supervised Learning: Generative or Contrastive.** *ArXiv, 1911.05722.*

Self-supervised contrastive learning
Uni-modal. SimCLR

- Sample pairs → data **augmentation** that doesn't change semantic information

- **Composition** of different image transformations: rotation, crop, gaussian blur, color distortion…

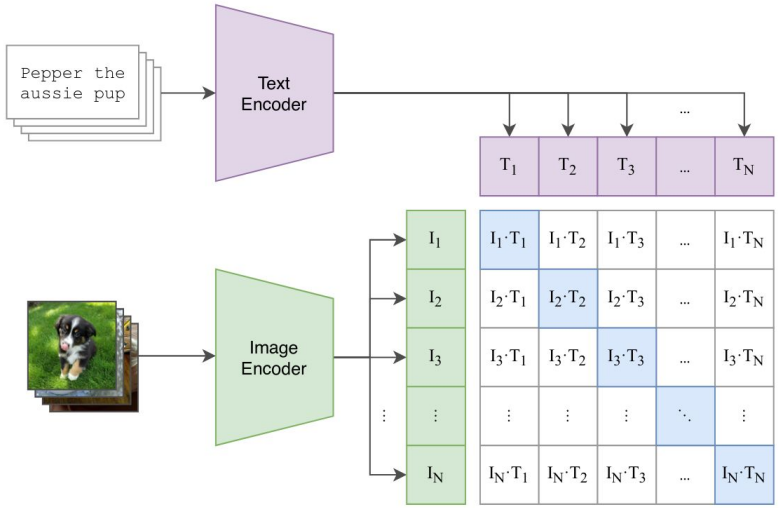


[original presentation]

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. (2020). **A simple framework for contrastive learning of visual representations.** In Daumé, H. and Singh, A., editors, Proceedings of the 37th International Conference on Machine Learning, volume 119 of Proceedings of Machine Learning Research (PMLR).

Self-supervised contrastive learning
Multi-modal. CLIP

- Natural **image and text** pairs. **400 million** pairs collected from the internet
- Impressive performance in image-to-text **zero-shot** classification
- Guided **image generation** (DALLE-2)
- **Foundation** model



Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I. (2021). **Learning transferable visual models from natural language supervision.** In Proceedings of the 38th International Conference on Machine Learning (ICML)

Self-supervised contrastive learning
Multi-modal. CLIP

- Natural **image and text** pairs. **400 million** pairs collected from the internet
- Impressive performance in image-to-text **zero-shot** classification
- Guided **image generation** (DALLE-2)
- **Foundation** model

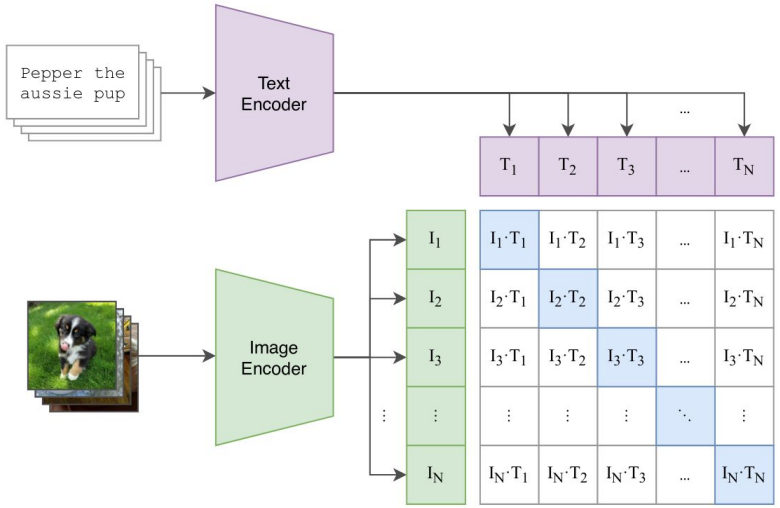
a shiba inu wearing a beret and black turtleneck

panda mad scientist mixing sparkling chemicals, artstation

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). **Learning transferable visual models from natural language supervision.** In Proceedings of the 38th International Conference on Machine Learning (ICML)
Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). **Hierarchical Text-Conditional Image Generation with CLIP Latents.** *ArXiv*.2204.06125

Self-supervised contrastive learning
Multi-modal. CLIP

- Natural **image and text** pairs. **400 million** pairs collected from the internet
- Impressive performance in image-to-text **zero-shot** classification
- Guided **image generation** (DALLE-2)
- **Foundation** model



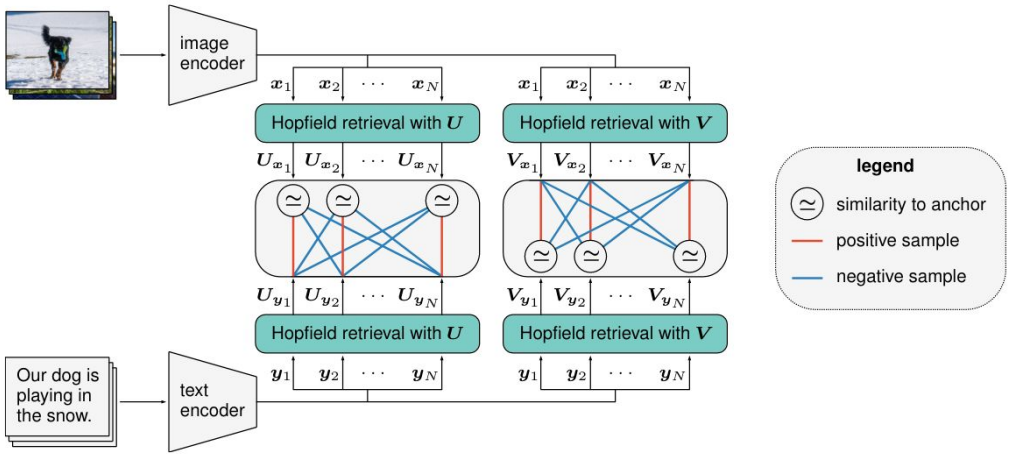
Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I. (2021). **Learning transferable visual models from natural language supervision.** In Proceedings of the 38th International Conference on Machine Learning (ICML)

Self-supervised contrastive learning
Multi-modal. CLOOB

- Continuous **modern Hopfield networks** extract co-occurrences and covariance structure
- **InfoLOOB** objective avoids saturation

$$\mathrm{L}_{\mathrm{InfoLOOB}} = -\ln \frac{\exp(\tau^{-1} \boldsymbol{x}_1^T \boldsymbol{y})}{\sum_{j=2}^{N} \exp(\tau^{-1} \boldsymbol{x}_j^T \boldsymbol{y})}$$

- **Outperforms CLIP** in zero-shot tasks



Fürst, A., Rumetshofer, E., Tran, V., Ramsauer, H., Tang, F., Lehner, J., David Kreil, Michael Kopp, Günter Klambauer, Angela Bitto-Nemling, Hochreiter, S. (2021). **CLOOB: Modern Hopfield Networks with InfoLoob outperform CLIP**. *ArXiv*. 2110.11316

Self-supervised contrastive learning
Multi-modal. CLOOB

- Continuous **modern Hopfield networks**

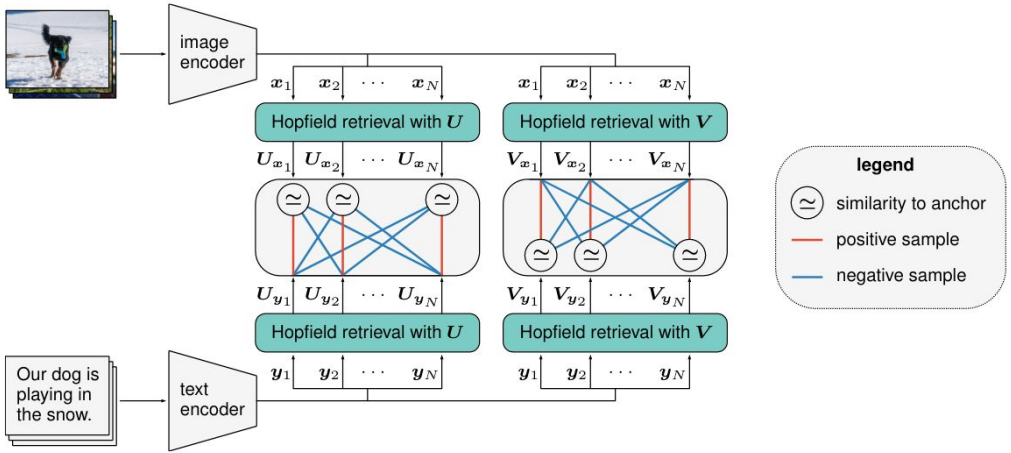
  extract co-occurrences and covariance

  structure

- **InfoLOOB** objective avoids saturation

$$\mathrm{L}_{\mathrm{InfoLOOB}} = -\ln \frac{\exp(\tau^{-1} \boldsymbol{x}_1^T \boldsymbol{y})}{\sum_{j=2}^{N} \exp(\tau^{-1} \boldsymbol{x}_j^T \boldsymbol{y})}$$

$$\mathrm{L}_{\mathrm{InfoNCE}} = -\ln \frac{\exp(\tau^{-1} \boldsymbol{x}_1^T \boldsymbol{y})}{\exp(\tau^{-1} \boldsymbol{x}_1^T \boldsymbol{y}) + \sum_{j=2}^{N} \exp(\tau^{-1} \boldsymbol{x}_j^T \boldsymbol{y})}$$



Fürst, A., Rumetshofer, E., Tran, V., Ramsauer, H., Tang, F., Lehner, J., David Kreil, Michael Kopp, Günter Klambauer, Angela Bitto-Nemling, Hochreiter, S. (2021). **CLOOB: Modern Hopfield Networks with InfoLoob outperform CLIP**. *ArXiv*. 2110.11316

Self-supervised contrastive learning
Multi-modal. CLOOB

- Continuous **modern Hopfield networks**

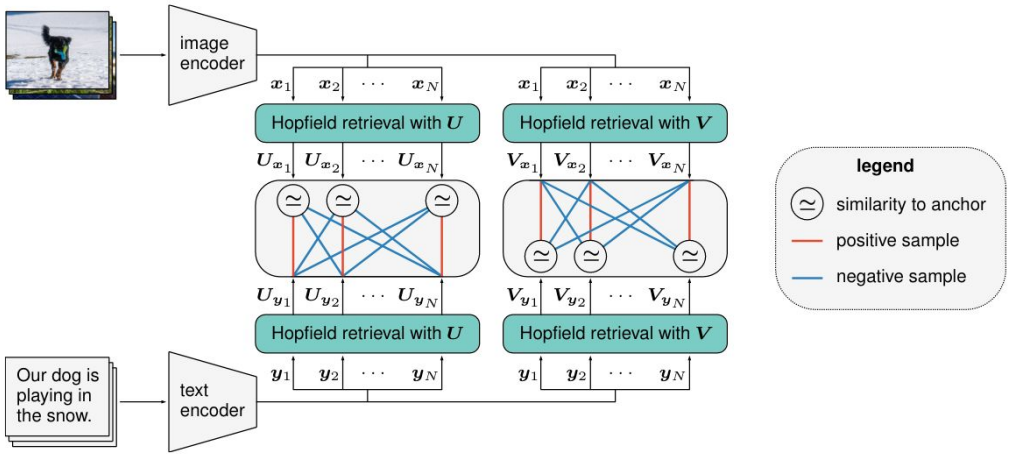
  extract co-occurrences and covariance

  structure

- **InfoLOOB** objective avoids saturation

$$\mathrm{L}_{\mathrm{InfoLOOB}} = -\ln \frac{\exp(\tau^{-1} \boldsymbol{x}_1^T \boldsymbol{y})}{\sum_{j=2}^{N} \exp(\tau^{-1} \boldsymbol{x}_j^T \boldsymbol{y})}$$

- **Outperforms CLIP** in zero-shot tasks



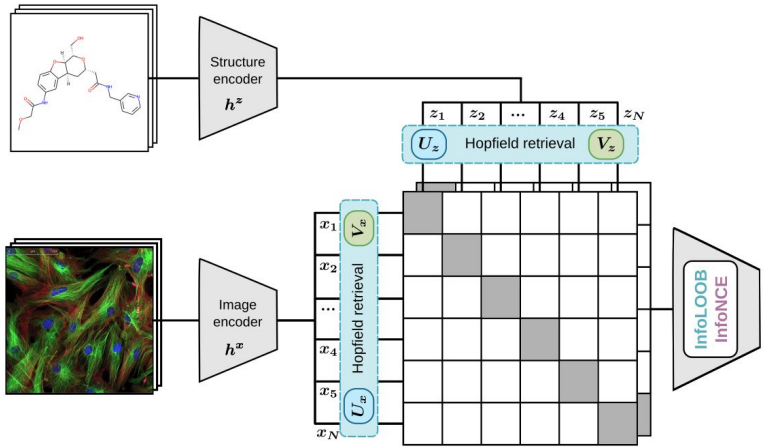
Fürst, A., Rumetshofer, E., Tran, V., Ramsauer, H., Tang, F., Lehner, J., David Kreil, Michael Kopp, Günter Klambauer, Angela Bitto-Nemling, Hochreiter, S. (2021). **CLOOB: Modern Hopfield Networks with InfoLoob outperform CLIP**. *ArXiv*. 2110.11316

Method

CLOOME. Contrastive Leave One Out boost for Molecule Encoder

Learn molecular **representations** with **contrastive** learning using

microscopy images and molecular **structures**



Sanchez-Fernandez, A., Rumetshofer, E., Hochreiter, S. & Klambauer, G. (2022, March). **Contrastive Learning of Image-and Structure-Based Representations in Drug Discovery.** In ICLR 2022 Machine Learning for Drug Discovery.

CLOOME. Contrastive Leave One Out boost for Molecule Encoder

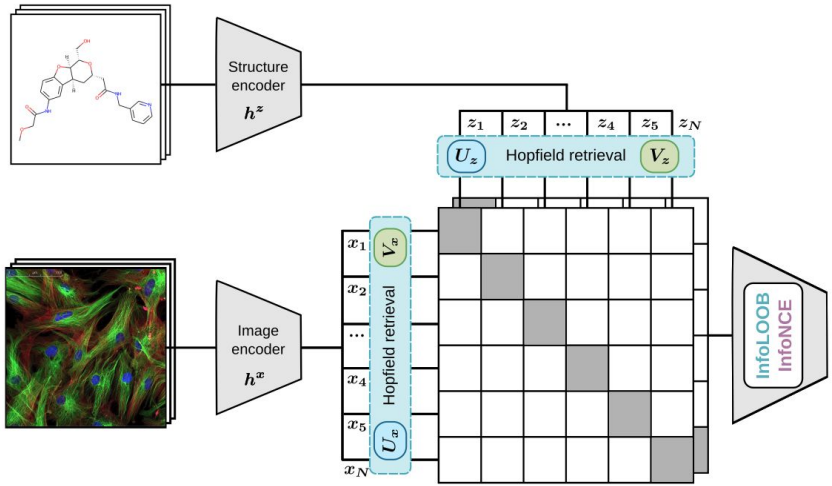


Image and structure pairs $\{(x_1, z_1), ..., (x_N, z_N)\}$

Embeddings

$$\boldsymbol{x_n} = \boldsymbol{h^x}(x_n) \qquad \boldsymbol{z_n} = \boldsymbol{h^z}(z_n)$$

$$\boldsymbol{X} = (\boldsymbol{x_1}, ..., \boldsymbol{x_N}) \qquad \boldsymbol{Z} = (\boldsymbol{z_1}, ..., \boldsymbol{z_N})$$

Modern Hopfield networks $\boldsymbol{U} = \boldsymbol{X}$ $\boldsymbol{V} = \boldsymbol{Z}$

Modern Hopfield network retrieved embeddings

$$\boldsymbol{U_{x_i}} = \boldsymbol{U}\text{softmax}(\beta \boldsymbol{U^T x_i}) \qquad \boldsymbol{V_{x_i}} = \boldsymbol{V}\text{softmax}(\beta \boldsymbol{V^T x_i})$$

$$\boldsymbol{U_{z_i}} = \boldsymbol{U}\text{softmax}(\beta \boldsymbol{U^T z_i}) \qquad \boldsymbol{V_{z_i}} = \boldsymbol{V}\text{softmax}(\beta \boldsymbol{V^T z_i})$$

$$\mathrm{L}_{\mathrm{InfoLOOB}} = -\frac{1}{N}\sum_{i=1}^{N} \ln \frac{\exp(\tau^{-1}\, \boldsymbol{U_{x_i}^T U_{z_i}})}{\sum_{j\neq i}^{N} \exp(\tau^{-1}\, \boldsymbol{U_{x_i}^T U_{z_j}})} - \frac{1}{N}\sum_{i=1}^{N} \ln \frac{\exp(\tau^{-1}\, \boldsymbol{V_{x_i}^T V_{z_i}})}{\sum_{j\neq i}^{N} \exp(\tau^{-1}\, \boldsymbol{V_{x_j}^T V_{z_i}})}$$

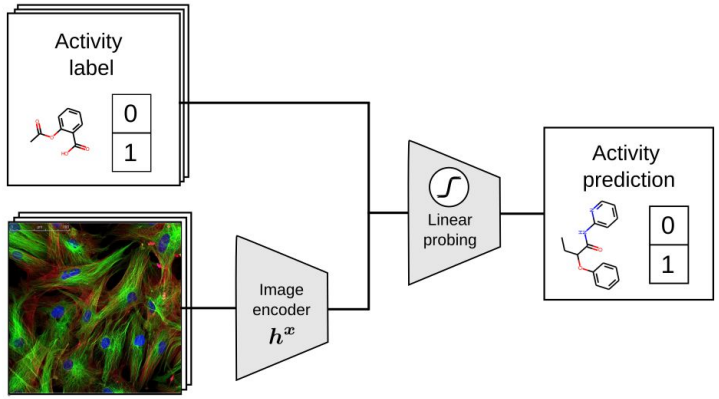
$$\mathrm{L}_{\mathrm{InfoNCE}} = -\frac{1}{N}\sum_{i=1}^{N} \ln \frac{\exp(\tau^{-1}\, \boldsymbol{x_i^T z_i})}{\sum_{j=1}^{N} \exp(\tau^{-1}\, \boldsymbol{x_i^T z_j})} - \frac{1}{N}\sum_{i=1}^{N} \ln \frac{\exp(\tau^{-1}\, \boldsymbol{x_i^T z_i})}{\sum_{j=1}^{N} \exp(\tau^{-1}\, \boldsymbol{x_j^T z_i})}$$

CLOOME. Contrastive Leave One Out boost for Molecule Encoder

- **Baselines**. CNNs and profile-based methods
- **Data split.**
  - Molecule and views from same sample in same split
  - Same as in Hofmarcher et al. + unlabeled samples added to training set
- **Batch size:** 256
- **Image downsizing:** 520x696 → 320x320

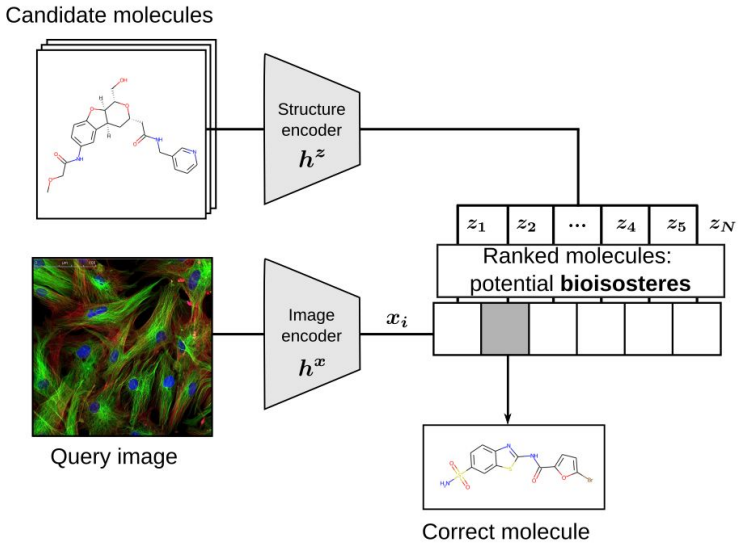
Results

Linear probing for bioactivity prediction



Type	Method	AUC	F1	AUC >0.9	AUC >0.8	AUC >0.7
Linear probing on self-supervised	CLOOME	**0.714**±0.20	0.395±0.32	57	84	109
Supervised	ResNet	**0.731**±0.19	0.508±0.30	68	94	119
	DenseNet	0.730±0.19	0.530±0.30	61	98	121
	GapNet	0.725±0.19	0.510±0.29	63	94	117
	MIL-Net	0.711±0.18	0.445±0.32	61	81	105
	M-CNN	0.705±0.19	0.482±0.31	57	78	105
	SC-CNN	0.705±0.20	0.362±0.29	61	83	109
	FNN	0.675±0.20	0.361±0.31	55	71	90

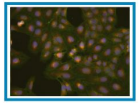
Retrieval for bioisosteric replacement

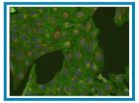


Method	Top-k accuracy (%)					
	Top-1	95% CI	Top-5	95% CI	Top-10	95% CI
CLOOME	**3.215**	[2.505, 4.058]	**6.998**	[5.947, 8.170]	**8.936**	[7.754, 10.233]
Random	0.047	[0.001, 0.263]	0.236	[0.077, 0.551]	0.473	[0.227, 0.868]

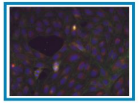
Retrieval for bioisosteric replacement

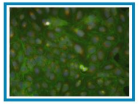
Query image



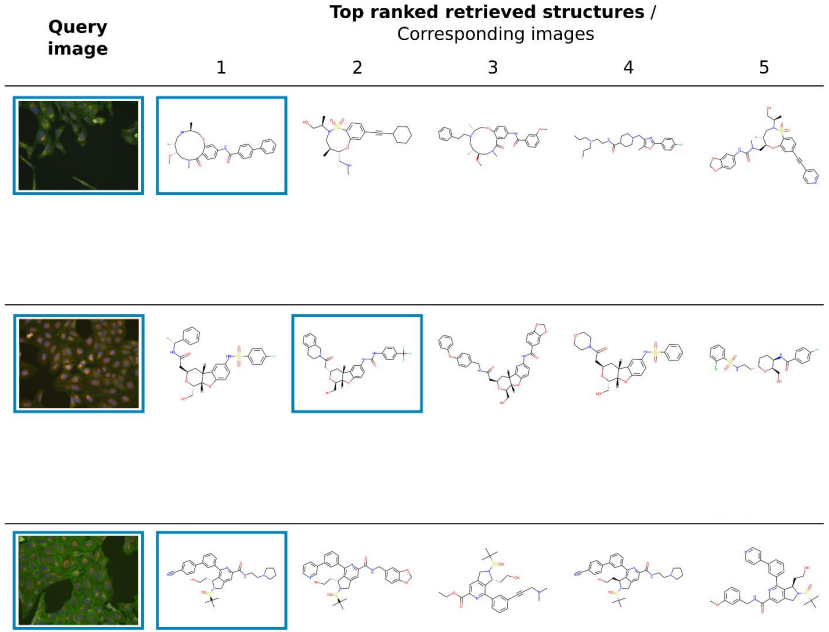


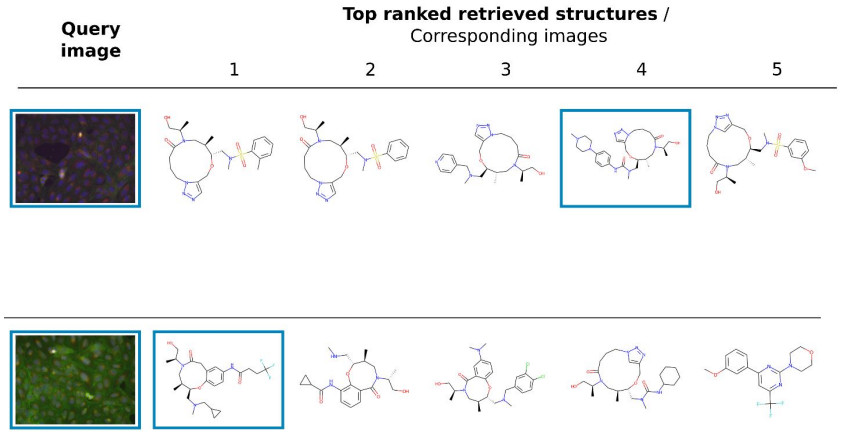
Query image



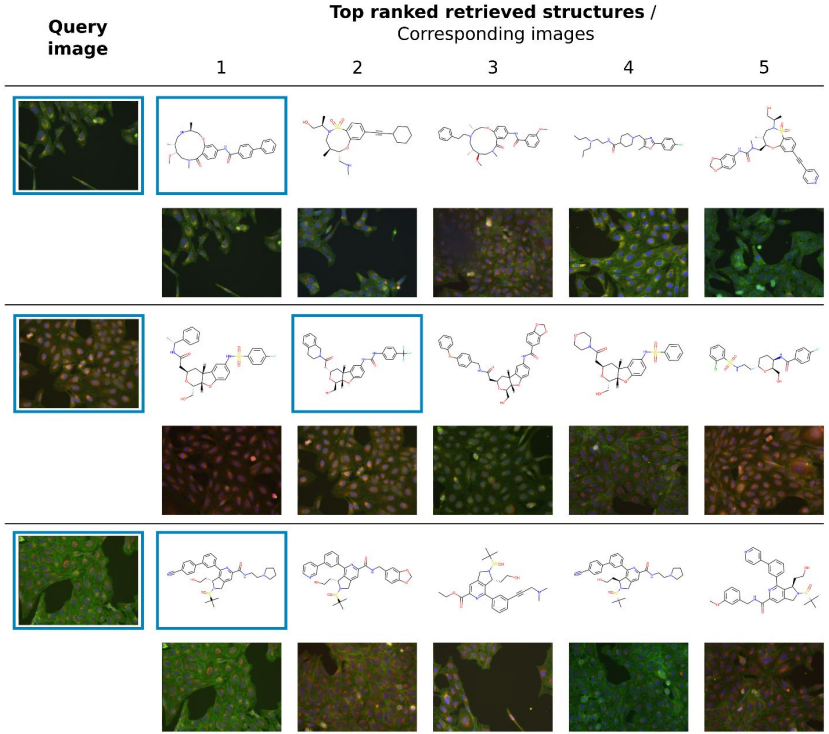


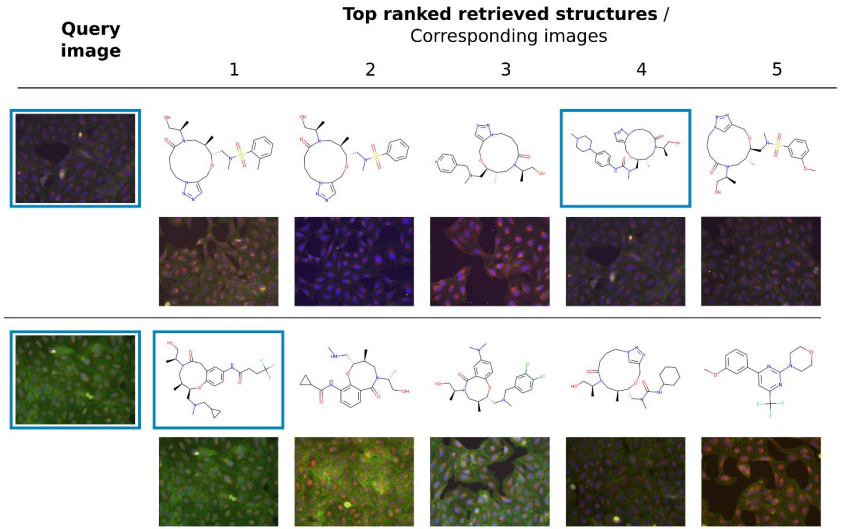
Retrieval for bioisosteric replacement





Retrieval for bioisosteric replacement





Conclusions

- **Microscopy images** could be useful for accelerating drug discovery
- **Contrastive learning** of **microscopy images** and chemical **structures** can produce **transferable representations** of molecules
- Including molecular structures and microscopy images could allow for **bioisosteric replacement** tasks

Next steps

- **Mechanism of Action (MoA)** as a downstream task
- **Analyze** obtained **representations**
- **JUMP-CP** dataset
  - 120,000 compounds
  - 10 pharma partners

Acknowledgements

- **Collaborators:** Elisabeth Rumetshofer, Günter Klambauer, Sepp Hochreiter

- **Institute for Machine Learning** previous work:

  - CNNs for microscopy.

    Hofmarcher, M., Rumetshofer, E., Clevert, D., Hochreiter, S., & Klambauer, G. (2019)**. Accurate Prediction of Biological Assays with High-Throughput Microscopy Images and Convolutional Networks.** *Journal Of Chemical Information And Modeling*, *59*(3), 1163-1171. doi: 10.1021/acs.jcim.8b00670

  - CLOOB.

    Fürst, A., Rumetshofer, E., Tran, V., Ramsauer, H., Tang, F., Lehner, J., David Kreil, Michael Kopp, Günter Klambauer, Angela Bitto-Nemling & Hochreiter, S. (2021). **CLOOB: Modern Hopfield Networks with InfoLoob outperform CLIP**. *ArXiv*. 2110.11316

- **AIDD** EU Horizon2020 project

References

Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). **Principles of early drug discovery.** *British Journal Of Pharmacology, 162(6)*, 1239-1249. doi: 10.1111/j.1476-5381.2010.01127.x

Cumming, J. G., Davis, A. M., Muresan, S., Haeberlein, M., & Chen, H. (2013). **Chemical predictive modelling to improve compound quality.** *Nature Reviews Drug Discovery*, *12*(12), 948–962. https://doi.org/10.1038/NRD4128

Schaduangrat, N., Lampa, S., Simeon, S., Gleeson, M. P., Spjuth, O., & Nantasenamat, C. (2020). **Towards reproducible computational drug discovery.** *Journal of Cheminformatics 2020 12:1*, *12*(1), 1–30. https://doi.org/10.1186/S13321-020-0408-X

Niranj Chandrasekaran, S., Ceulemans, H., Boyd, J. D., & Carpenter, A. E. (2021). **Image-based profiling for drug discovery: due for a machine-learning upgrade?** *Nature Reviews Drug Discovery*. https://doi.org/10.1038/s41573-020-00117-w

Lin, S., Schorpp, K., Rothenaigner, I., & Hadian, K. (2020). **Image-based high-content screening in drug discovery.** *Drug Discovery Today, 25(8)*, 1348-1361. doi: 10.1016/j.drudis.2020.06.001

Bray, M. A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., Kost-Alimova, M., Gustafsdottir, S. M., Gibson, C. C., & Carpenter, A. E. (2016). **Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes.** *Nature Protocols 2016 11:9*, *11*(9), 1757–1774. https://doi.org/10.1038/nprot.2016.105

Simm, J., Klambauer, G., Arany, A., Steijaert, M., Wegner, J. K., Gustin, E., Chupakhin, V., Chong, Y. T., Vialard, J., Buijnsters, P., Velterm Ingrid, Vapirev, A., Singh, S., Carpenter, A. E., Wuyts, R., Hochreiter, S., Moreau, Y., & Ceulemans, H. (2018). **Repurposing high-throughput image assays enables biological activity prediction for drug discovery.** *Cell Chem Biol*, *25*(5), 611–618. https://doi.org/10.1016/j.chembiol.2018.01.015

Hofmarcher, M., Rumetshofer, E., Clevert, D., Hochreiter, S., & Klambauer, G. (2019)**. Accurate Prediction of Biological Assays with High-Throughput Microscopy Images and Convolutional Networks.** *Journal Of Chemical Information And Modeling*, *59*(3), 1163-1171. doi: 10.1021/acs.jcim.8b00670

Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., Tang, J., & Fellow, I. (2021) **Self-supervised Learning: Generative or Contrastive.** ArXiv, 1911.05722.

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). **A simple framework for contrastive learning of visual representations.** In Daumé, H. and Singh, A., editors, Proceedings of the 37th International Conference on Machine Learning, volume 119 of Proceedings of Machine Learning Research (PMLR), pages 1597–1607

K. He, H. Fan, Y. Wu, S. Xie & R. Girshick. (2020). **Momentum Contrast for Unsupervised Visual Representation Learning.** 2020 IEEE/CVF *Conference on Computer Vision and Pattern Recognition* (CVPR), 2020, pp. 9726-9735, doi: 10.1109/CVPR42600.2020.00975.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). **Learning transferable visual models from natural language supervision.** In *Proceedings of the 38th International Conference on Machine Learning* (ICML)

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). **Hierarchical Text-Conditional Image Generation with CLIP Latents.** ArXiv.2204.06125

Fürst, A., Rumetshofer, E., Tran, V., Ramsauer, H., Tang, F., Lehner, J., David Kreil, Michael Kopp, Günter Klambauer, Angela Bitto-Nemling & Hochreiter, S. (2021). **CLOOB: Modern Hopfield Networks with InfoLoob outperform CLIP**. *ArXiv*. 2110.11316

Thank you!
Questions?