Deep-learning in-situ classification of HIV-1 virion morphology

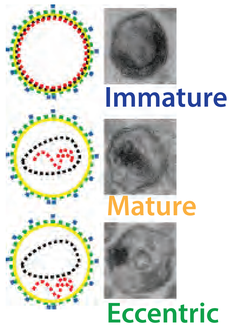
Juan S. Rey[1], Wen Li[2], Alexander J. Bryer[1], Hagan Beatson[1], Christian Lantz[1], Alan N. Engelman[2], Juan R. Perilla[1]
[1]Department of Chemistry and Biochemistry, University of Delaware
[2]Department of Medicine, Harvard University

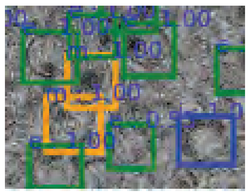
TEM on HIV discovery

Transmission Electron Microscopy (TEM) is critical for studying HIV maturation inhibitors.

- Removal of IN domain of Gag-Pol increases the frequency of immature particles.
- IN missense mutations can elicit eccentric virus formation where the RNP appears outside the viral capsid.

Impeding viral maturation leads to noninfectous virus.

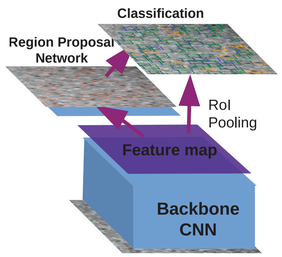




Assessing the efficacy of an inhibitor requires manual counting and classification of viral morphologies from TEM samples.

We can use Deep Learning to automate this process!

TEMNet Architecture



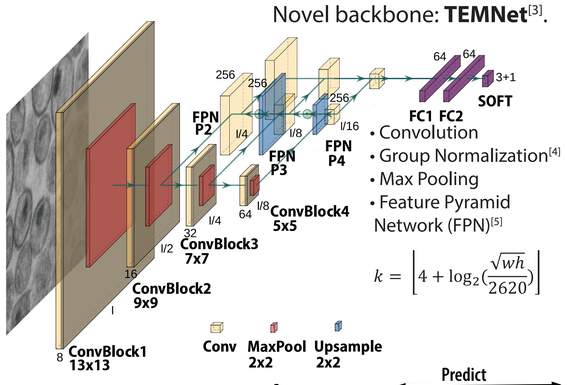
Faster RCNN[2] two-stage multiple object detection architecture.

$$L(\{\hat{p}_i, c_i\}, \{\hat{t}_i, t_i\}) = \frac{1}{N_{class}} \sum_i L_{class}(\hat{p}_i, c_i) + \frac{\lambda}{N_{RoI}} \sum_i L_{loc}(\hat{t}_i, t_i)$$

$$L_{class}(\hat{p}_i, c_i) = -\log\left(\frac{e^{\hat{p}_{i,c_i}}}{\sum_n^C e^{\hat{p}_{i,n}}}\right)$$

$$L_{loc}(\hat{t}_i, t_i) = \sum_{r \in \{x,y,w,h\}} \text{smooth}L1(\hat{t}_{i,r} - t_{i,r})$$

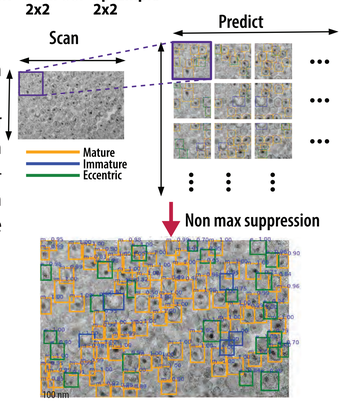
Novel backbone: **TEMNet**[3].



- Convolution
- Group Normalization[4]
- Max Pooling
- Feature Pyramid Network (FPN)[5]

$$k = \left\lfloor 4 + \log_2\left(\frac{\sqrt{wh}}{2620}\right) \right\rfloor$$

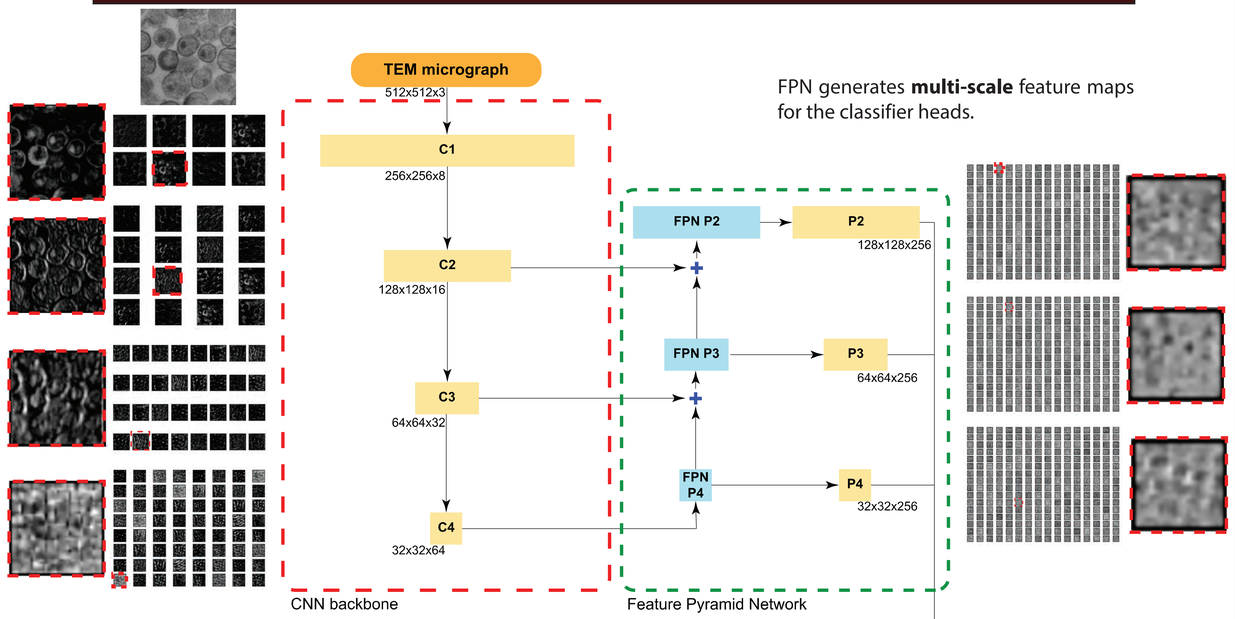
Sliding window approach to generate predictions. Window resizes to fit for different TEM micrograph magnifications and generates predictions on magnifications that were not used for training.



$$r = W_t / M_t [px]$$

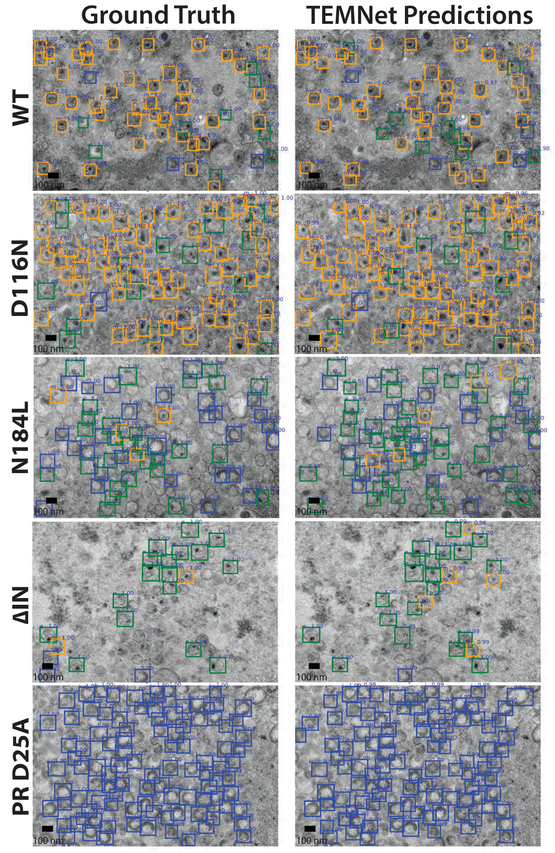
$$W_{new} = r\, M_{new} [px]$$

TEMNet: Deep learning for TEM classification

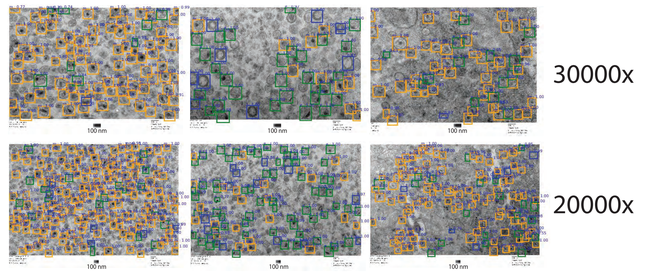


First convolutional block identifies edges and electron density in virion lumina and second block learns image processing watershedding.

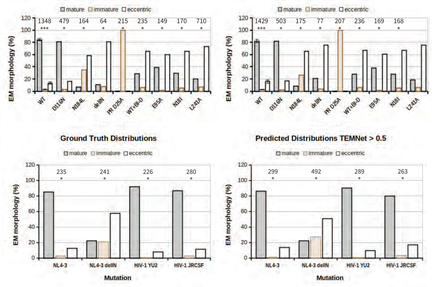
Performance on mutant samples



Multi-magnification predictions



Predicted morphology distributions are **statistically significant** compared to manually ascribed ground truth distributions for in situ micrographs across different samples including WT, IN and PR mutant viruses and primary isolates.

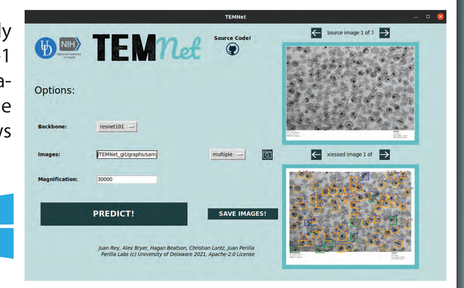


RCNN backbone	$mAP_{IoU=0.5}$%	Precision %	Recall %	F1 score %
ResNet101 + FPN	78.2	76.7	84.0	80.0
ResNet101v2 + FPN	77.9	75.8	85.9	80.5
InceptionResNetv2 + FPN	78.6	64.7	86.2	73.7
TEMNet (BN)	77.8	60.8	84.9	70.9
TEMNet (GN)	80.0	71.3	87.9	78.8

TEMNet achieves an **80.0%** mAP across a diverse set of experimental samples and primary isolates, outperforming other common CNNs.

TEMNet GUI

User friendly GUI for HIV-1 virion classification is available for Windows and Linux!



Conclusions

We have developed an end-to-end deep learning solution to the automated detection and classification of HIV-1 virion particle morphologies from TEM micrographs across different maturation stages.

Our methodology provides advantages to virologists and medical researchers:

- Accuracy across a diverse set of experimental samples including primary isolates.
- Is extremely efficient, increasing the speed in data processing by 3 orders of magnitude (7800 viral particles classified in 5 minutes).
- Our network overcomes the limitations of small datasets due to data augmentation strategies.
- TEMNet is light, only 15MB in size makes it appropriate for deployment of web and mobile applications.
- Can be extended to other enveloped viruses (e.g. SARS-CoV-2) provided enough training data for the machine learning algorithm.

We hope our method proves useful in antiretroviral drug development based on Gag-interacting maturation inhibitors and ALLINIs, which are currently in preclinical development.

All our code and application is freely available!
Get it here:

https://github.com/Perilla-lab/TEMNet

References

1. Sundquist, W. I., & Kräusslich, H. G. (2012). HIV-1 assembly, budding, and maturation. Cold Spring Harbor perspectives in medicine, 2(7), a006924.
2. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.
3. Rey, J. S., Li, W., Bryer, A. J., Beatson, H., Lantz, C., Engelman, A. N., & Perilla, J. R. (2021). Deep-learning in situ classification of HIV-1 virion morphology. Computational and structural biotechnology journal, 19, 5688–5700.
4. Wu, Y., & He, K. (2018). Group normalization. In Proceedings of the European conference on computer vision (ECCV) (pp. 3-19).
5. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).

Acknowledgements

This work was supported by NSF grants MCB-2027096, OAC-1818253 and ACI-1548562.