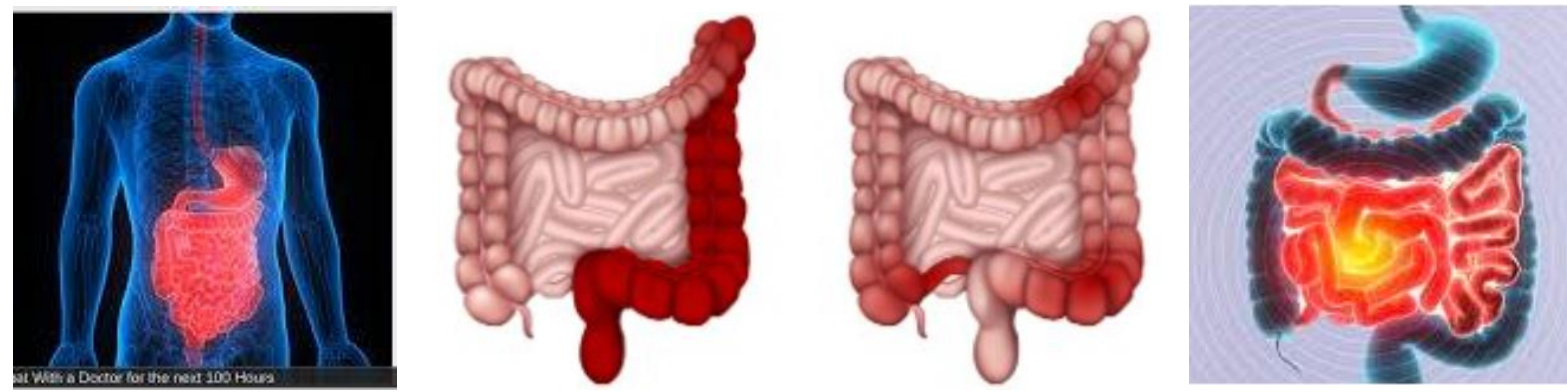


# DEEP LEARNING ALGORITHMS DERIVED FROM WEARABLE DEVICE METRICS PREDICT FUTURE INFLAMMATORY BOWEL DISEASE FLARES

Icahn School of Medicine at Mount Sinai



Kyung Won Lee<sup>1</sup>, Matteo Danieletto<sup>2</sup>, Jessica K. Whang<sup>3</sup>, Kyle Landell<sup>2</sup>, Drew Helmus<sup>3</sup>, Mayte Suarez-Farinas<sup>1</sup>, Bruce E. Sands<sup>3</sup>, Robert P. Hirten<sup>3</sup>

1. Department of Population Health Science and Policy  
2. The Hasso Plattner Institute for Digital Health  
Icahn School of Medicine at Mount Sinai, New York, NY  
3. Division of Gastroenterology



## INTRODUCTION

- Inflammatory bowel disease (IBD) is a chronic inflammatory disease of the gastrointestinal tract. Flares are common and unpredictable, presenting significant challenges for proactive disease management.
- Current monitoring approaches, including blood and stool biomarkers, imaging, and colonoscopy, are invasive, episodic, and typically employed only after symptoms emerge.
- Wearable devices offer a non-invasive, passive means of capturing continuous physiological signals that may correlate with underlying IBD inflammation.<sup>1-3</sup> (i.e., COVID-19<sup>4</sup>)

## AIM

- To develop wearable-based signatures, leveraging deep learning models and to predict IBD flares before they arise.

## METHODS

- The IBD Forecast study is a prospective cohort study enrolling adults ( $\geq 18$  years of age) in the United States with IBD who is willing to:
  - (1) Use a commercially available wearable device (Apple Watch)
  - (2) Download our custom ehive app



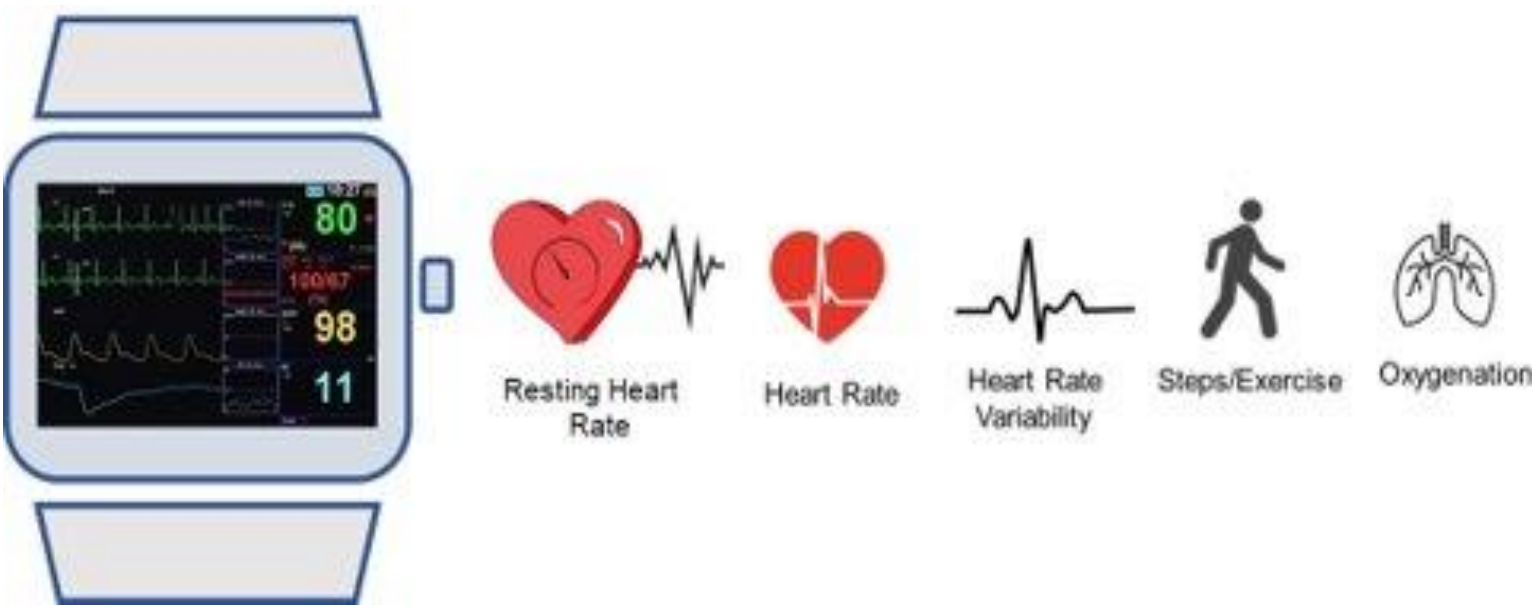
(3) Complete daily symptom assessments using Patient-Reported Outcome-2 (PRO-2)



(4) Followed for up to 18 months.

### INPUT Definition: Predictors (X)

- Apple Watch metrics, such as heart rate (HR), heart rate variability (HRV), and oxygen saturation (SpO<sub>2</sub>), were collected and included in the deep learning model.



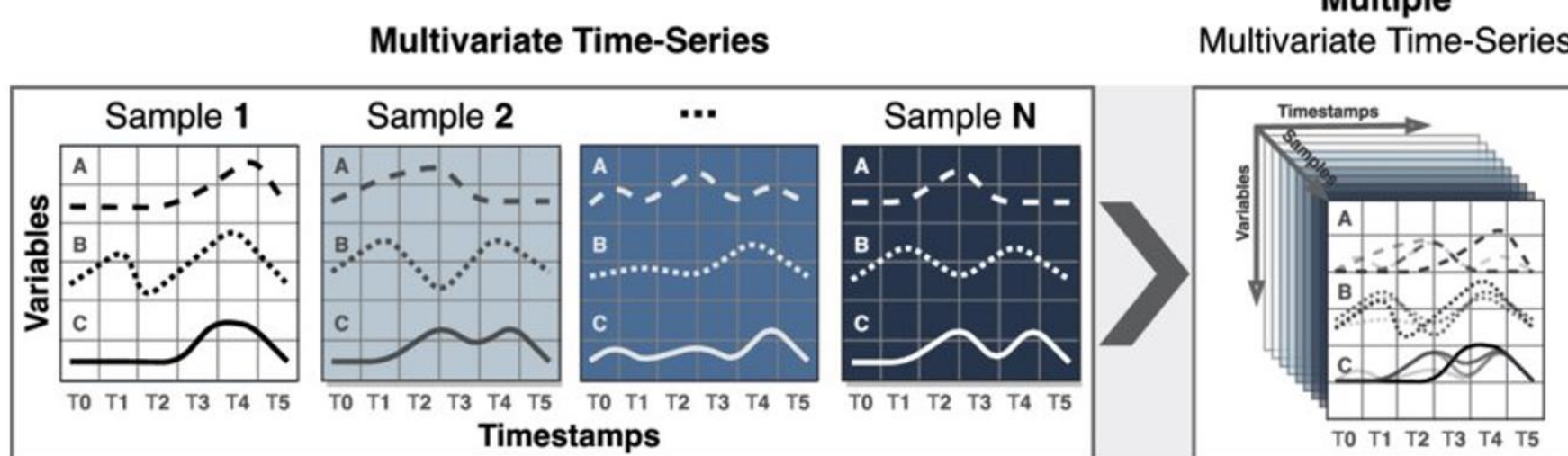
## METHODS

### OUTCOME Definition (Y)

- Symptom status was defined by daily PRO-2 surveys (asymptomatic; PRO-2 CD  $< 8$ ; PRO-2 UC  $\leq 1$  with rectal bleeding score = 0, stool frequency score  $\leq 1$ ).
- A symptomatic flare was defined as  $\geq 2$  symptomatic days in a 7-day period.
- Inflammatory flares were defined as a C-reactive protein  $> 5\text{mg/L}$ , sedimentation rate  $> 30\text{mm/hr}$ , or a fecal calprotectin  $> 150\mu\text{g/g}$ .

### MODELS

- Deep learning algorithms, including long short-term memory (LSTM) networks<sup>5</sup> and Transformer models<sup>6</sup>, were applied.



- The models were trained using binary focal cross-entropy loss<sup>7</sup> with the Adam optimizer<sup>8</sup>.
- The models assess the predictive utility of these physiological metrics for a composite endpoint of symptomatic or inflammatory flares up to 8 weeks before their onset.

## RESULTS

- A total of 140 participants were enrolled and followed for an average of 7 months.

Table 1. Baseline demographic information.

	Crohn's Disease (CD) (N=77)	Ulcerative Colitis (UC) (N=63)	Overall (N=140)
Age, years (SD)	39.2 (13.7)	39.9 (12.6)	39.5 (13.2)
Sex, female (%)	51 (66.2)	36 (57.1)	87 (62.1)
Race (%)			
Asian	2 (2.6)	5 (7.9)	7 (5.0)
Black	3 (3.9)	2 (3.2)	5 (3.6)
Native-American	1 (1.3)	1 (1.6)	2 (1.4)
White	69 (89.6)	54 (85.7)	123 (87.9)
Ethnicity (%)			
Hispanic	5 (6.5)	4 (6.3)	9 (6.4)
Not Hispanic	72 (93.5)	57 (90.5)	129 (92.1)
Smoking (%)			
Current	2 (2.6)	2 (3.2)	4 (2.9)
Never	61 (79.2)	45 (71.4)	106 (75.7)
Past	14 (18.2)	16 (25.4)	30 (21.4)

- Using the physiological metrics from the Apple Watch, both the Transformer model (AUC 0.83; Sensitivity 0.78; Specificity 0.68; AUPRC 0.31; F1-score 0.28) and LSTM model (AUC 0.82; Sensitivity 0.80; Specificity 0.66; AUPRC 0.33; F1-score 0.27) predicted inflammatory and symptomatic flares up to 8 weeks in advance.
- Peak AUC performance in the testing set for the Transformer model was observed 35 days before flares (AUC 0.83; Sensitivity 0.78; Specificity 0.68; AUPRC 0.31; F1-score 0.28).

## STUDY INFORMATION

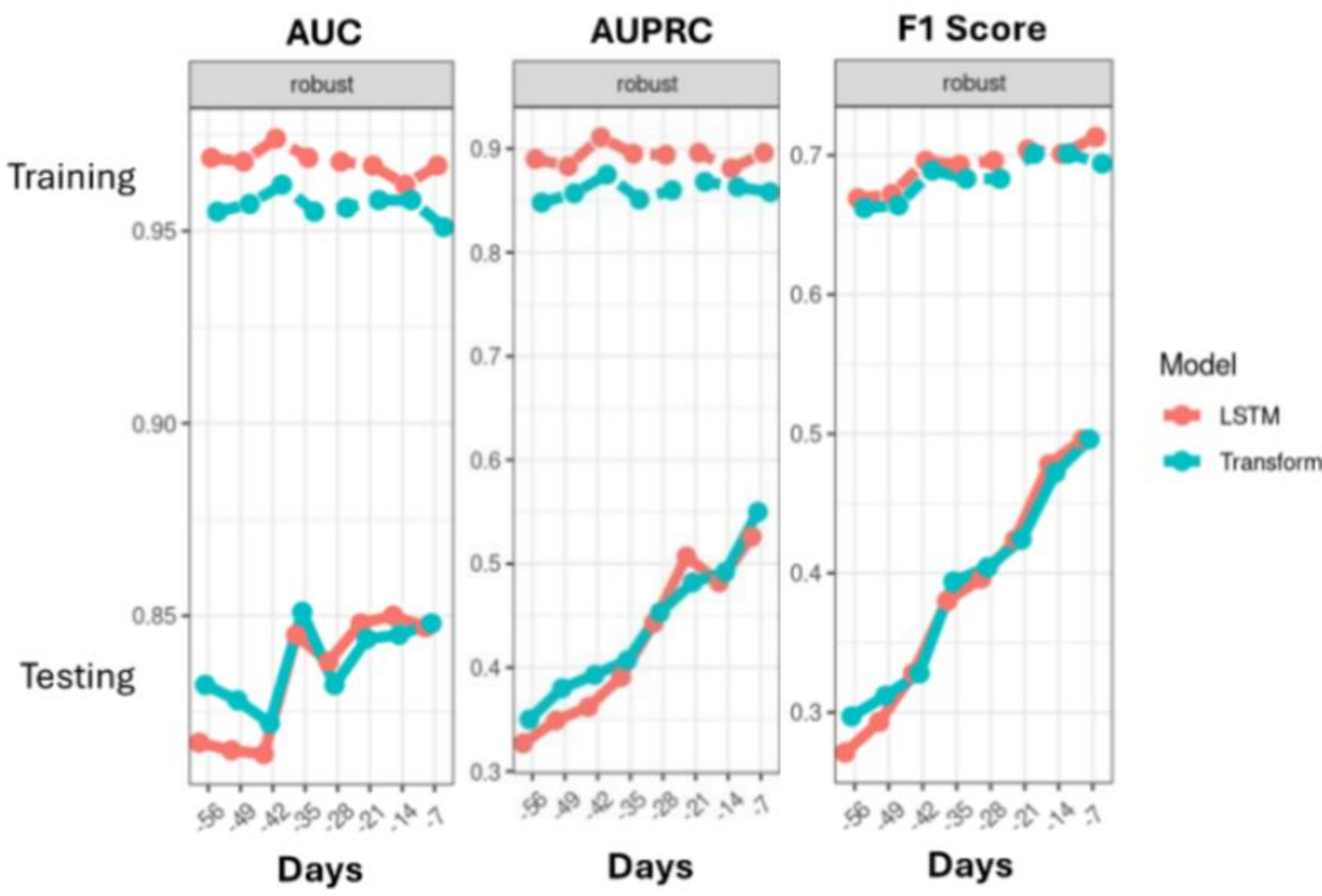


IBD Forecast  
A DIGITAL STUDY  
[www.hpiims.org/research-projects/ibd-forecast](http://www.hpiims.org/research-projects/ibd-forecast)

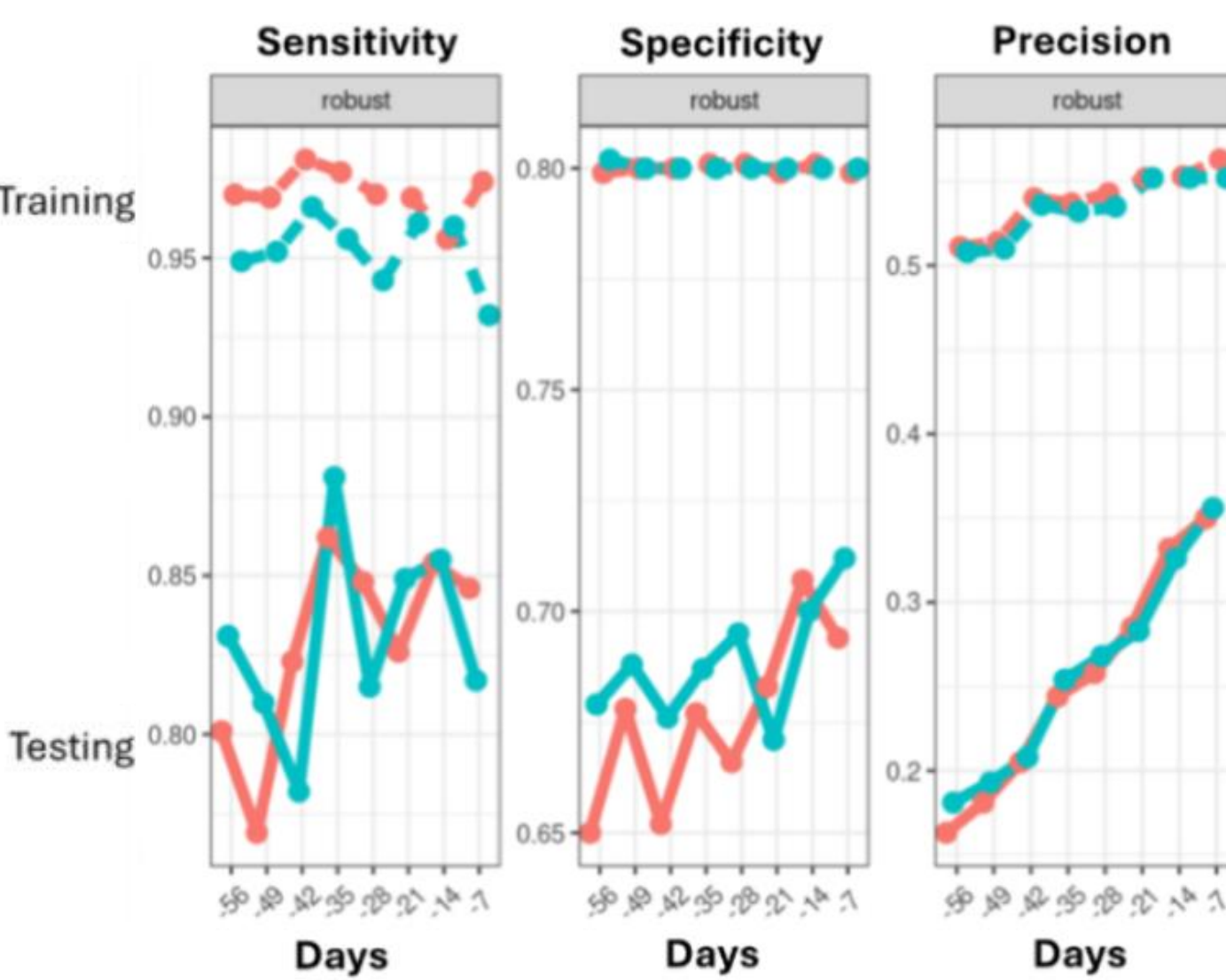


## RESULTS

**Figure 1.** Performance metrics for LSTM and Transformer deep learning models in predicting IBD flares up to 8 weeks before flare occurrence. Dashed lines denote the training set results, while solid lines denote the testing set results. Pink represents LSTM model and blue represents Transformer model.



- The models were established with a maximum sensitivity of  $\geq 0.90$  and specificity of  $\geq 0.80$  in the training set.



- Both models demonstrated strong predictive performance, with Transformer models outperforming LSTM models.

## CONCLUSIONS

- Longitudinally collected physiological metrics from wearable devices successfully predict IBD flares up to 8 weeks in advance.
- These findings highlight the potential of wearable devices to be used for proactive disease management.

## REFERENCES

- Hirten, Robert P., et al. "Physiological Data Collected From Wearable Devices Identify and Predict Inflammatory Bowel Disease Flares." *Gastroenterology* (2025).
- Hirten, Robert, et al. "Physiological metrics collected from wearable devices identify inflammatory and clinical inflammatory bowel disease flares." *Inflammatory Bowel Diseases* 29, Supplement\_1 (2023): S21-S22.
- Hirten, Robert P., et al. "Wearable devices are well accepted by patients in the study and management of inflammatory bowel disease: a survey study." *Digestive Diseases and Sciences* 66 (2021): 1836-1844.
- Hirten, Robert P., et al. "Use of physiological data from a wearable device to identify SARS-CoV-2 infection and symptoms and predict COVID-19 diagnosis: observational study." *Journal of medical Internet research* 23.2 (2021): e26107.
- Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.
- Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- Lin, Tsung-Yi, et al. "Focal loss for dense object detection." *Proceedings of the IEEE international conference on computer vision*. 2017.
- Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).