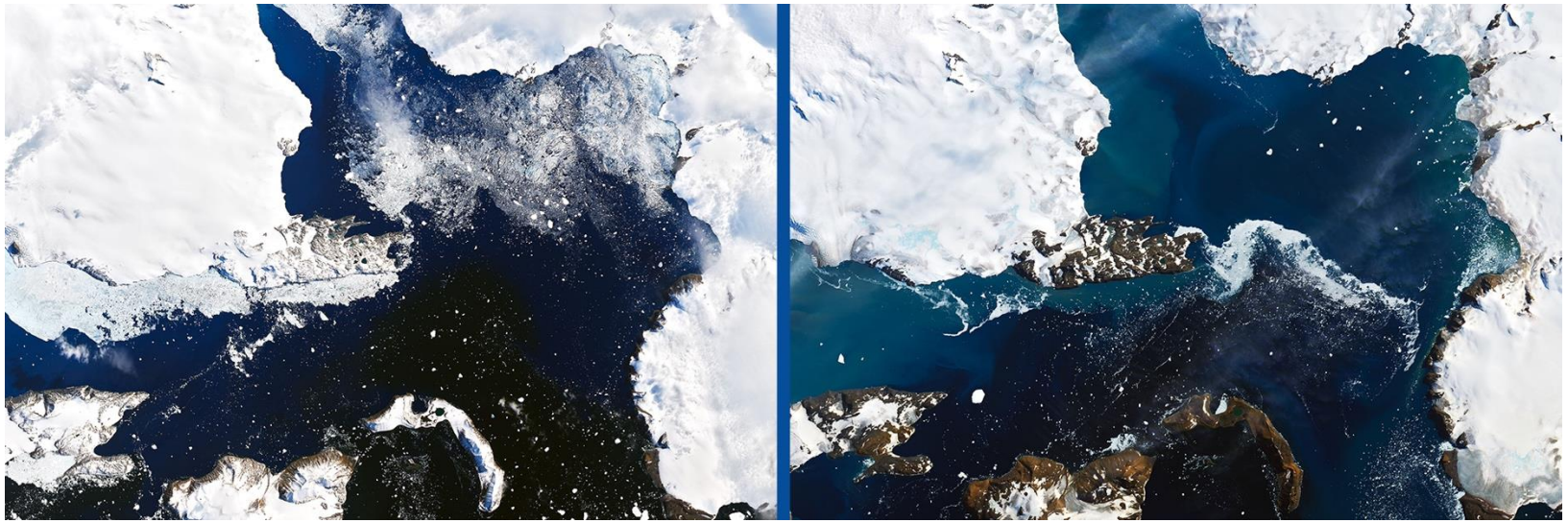
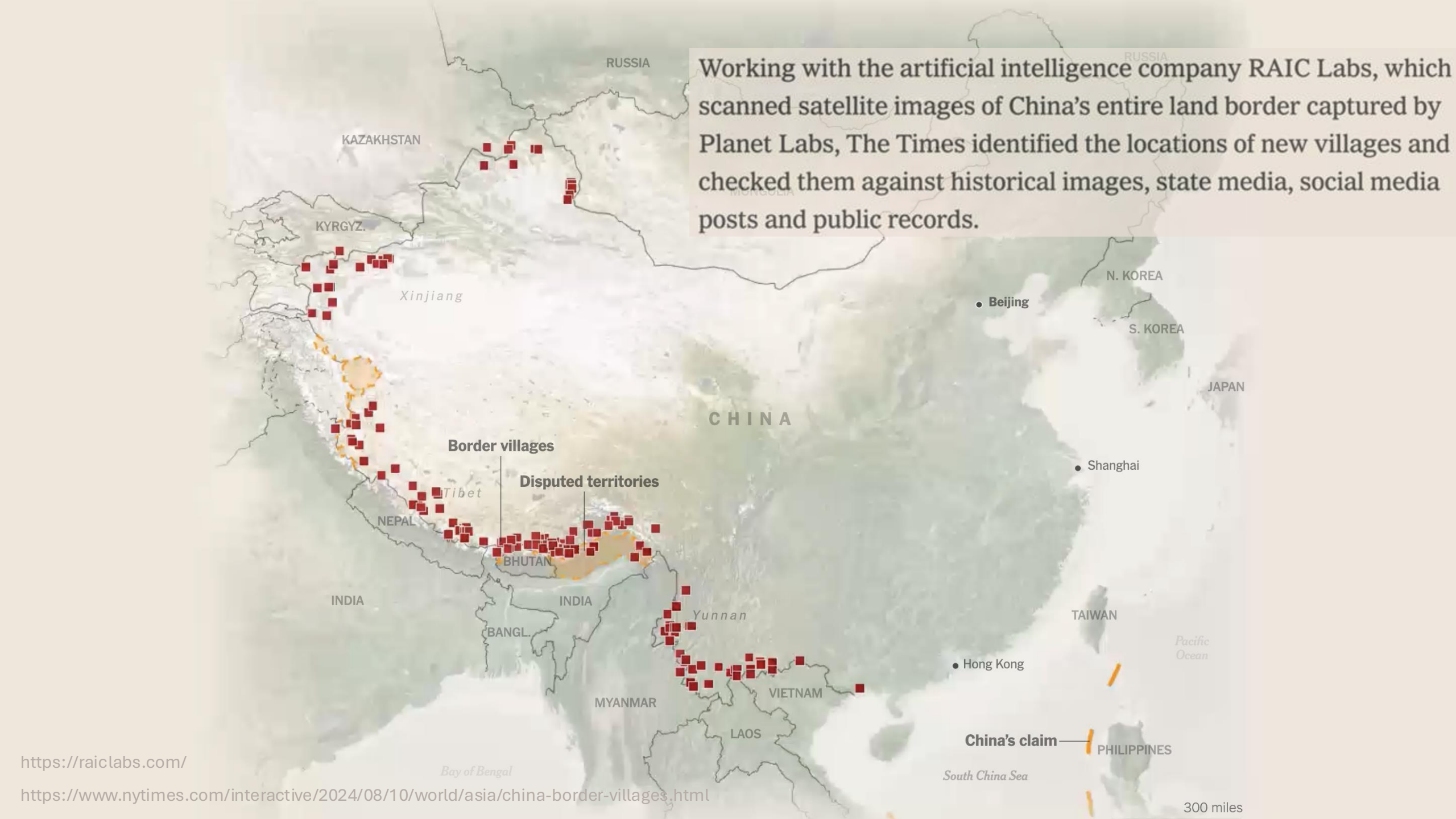
https://globalnews.ca/news/6593386/antartica-melting-nasa-photos/



https://theclickhub.com/how-google-is-helping-to-tackle-global-deforestation/

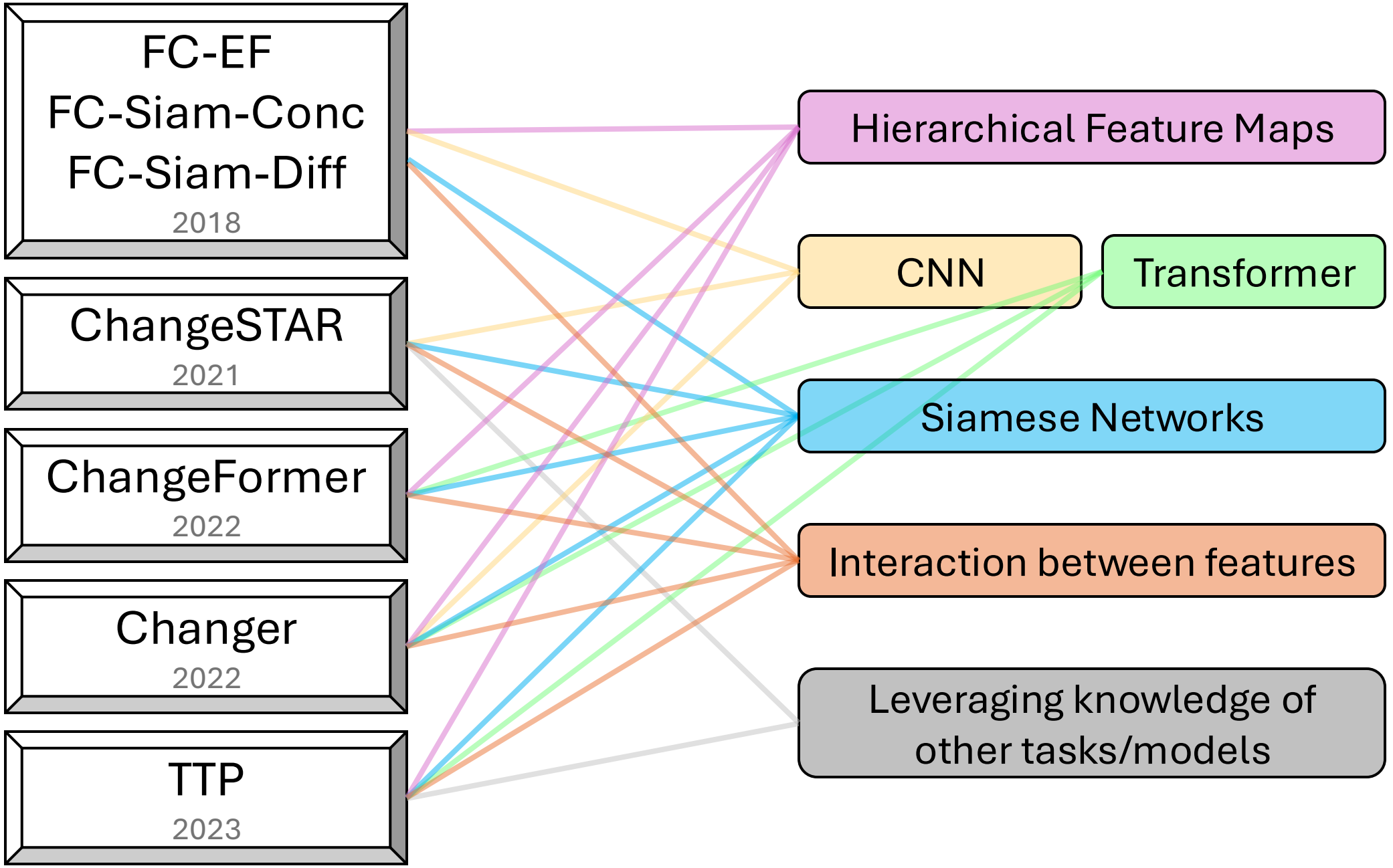
https://www.nytimes.com/interactive/2024/06/09/world/middleeast/gaza-tents.html



https://raiclabs.com/

https://www.nytimes.com/interactive/2024/08/10/world/asia/china-border-villages.html

Models & Characteristics



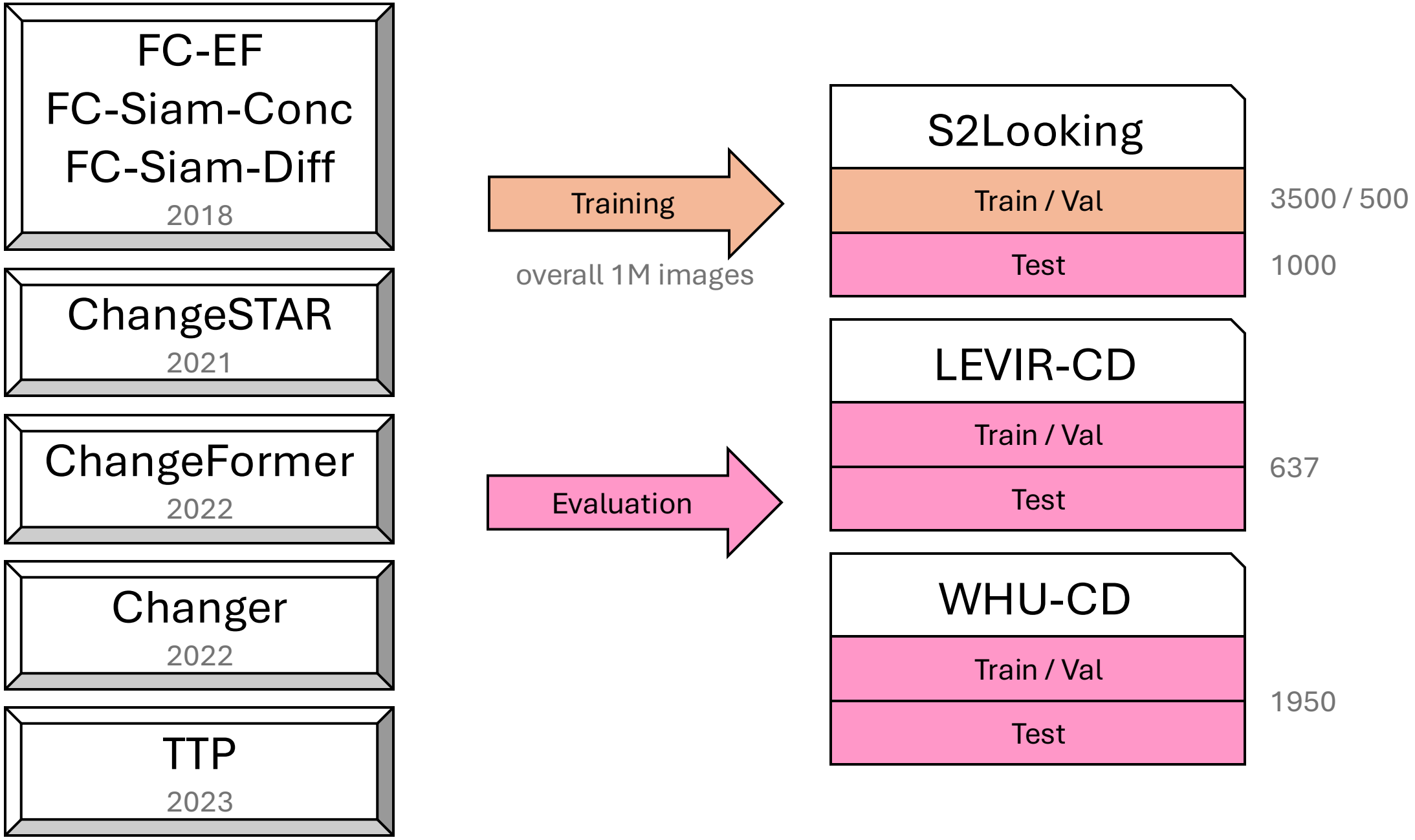
Metrics

Metric	Accuracy	Recall	Precision	F1 Score	IoU
Formula	$\frac{TP+TN}{TP+TN+FP+FN}$	$\frac{TP}{TP+FN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+\frac{FN+FP}{2}}$	$\frac{TP}{TP+FN+FP}$

		Predicted condition	
	Total population = P + N	**Positive (PP)** *(Change)*	**Negative (PN)** *(No change)*
Actual condition	**Positive (P)** *(Change)*	**True positive (TP)**	**False negative (FN)**
	Negative (N) *(No change)*	**False positive (FP)**	**True negative (TN)**

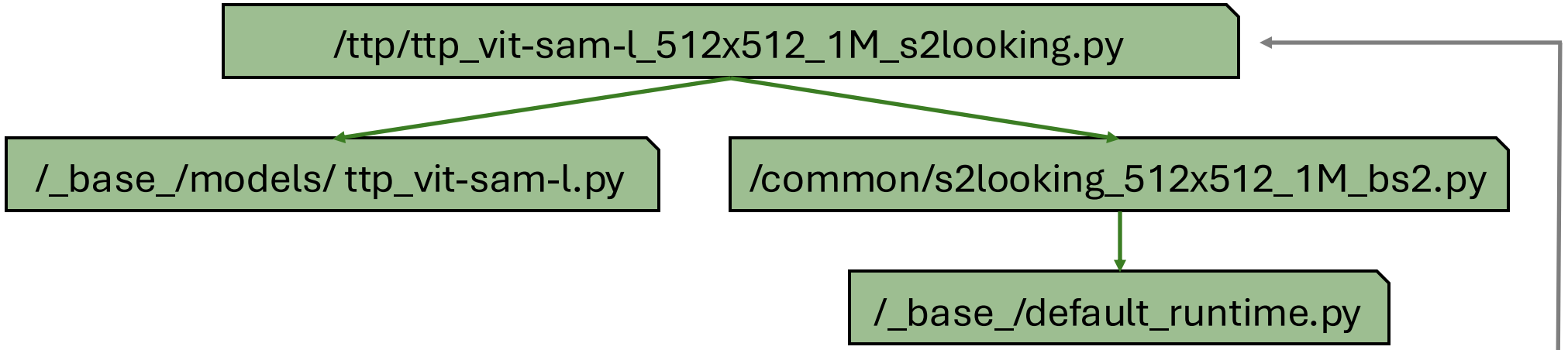
https://en.wikipedia.org/wiki/Confusion_matrix

Experimental Setup



Open-CD Framework

Configuration files determine model architecture and training specifics:



Code of framework not adapted, but executed with correct configuration file:

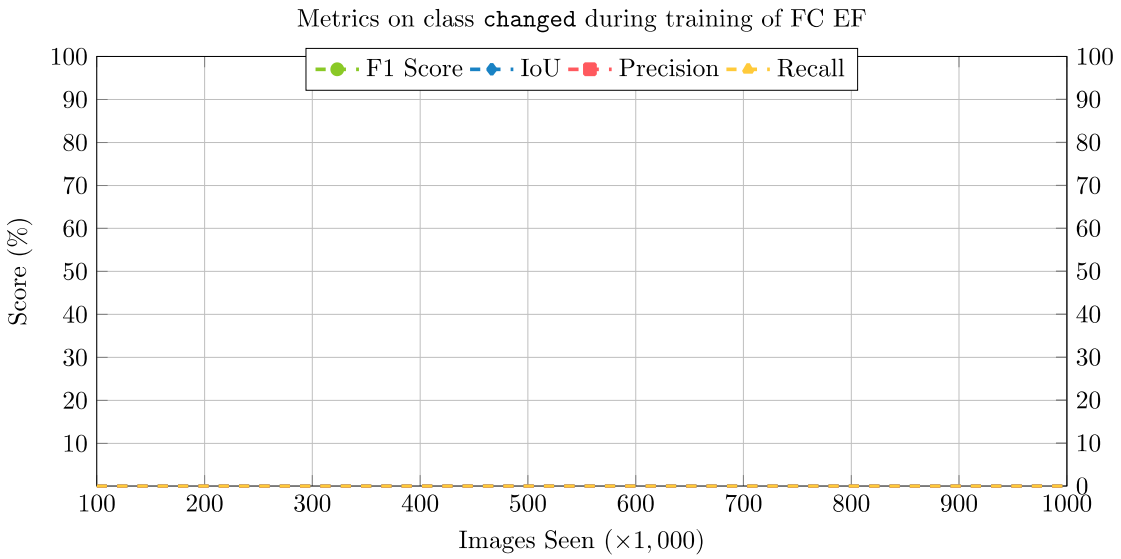
*$ python open-cd/tools/**train.py***
open-cd/configs/ttp/ttp_vit-sam-l_512x512_1M_s2looking.py
--work-dir ./ttp

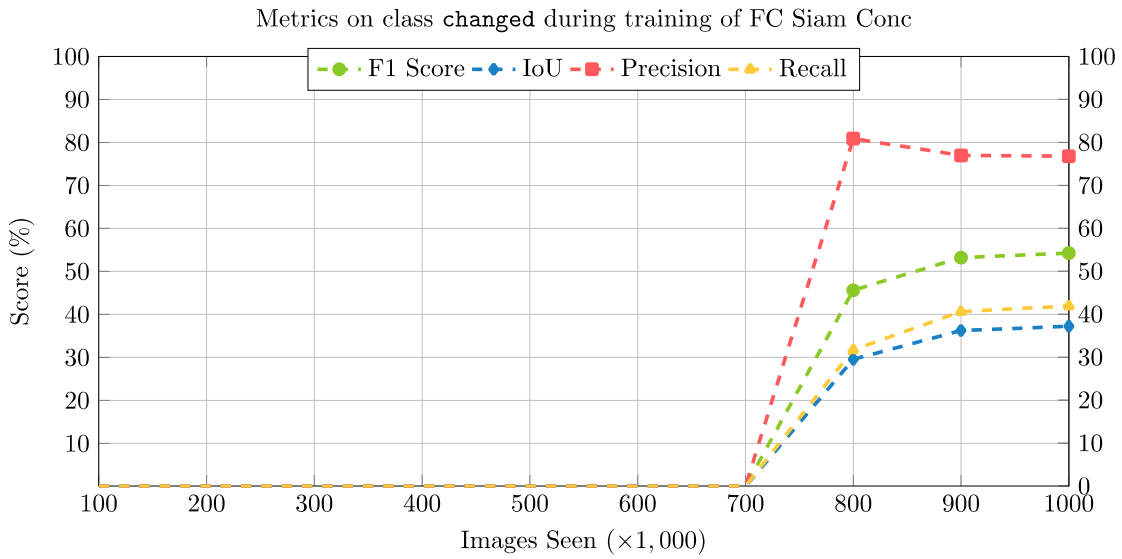
*$ python open-cd/tools/**test.py***
open-cd/configs/ttp/ttp_vit-sam-l_512x512_1M_s2looking.py
ttp/iter_325000.pth
--work-dir ./ttp --show-dir ./ttp/s2looking

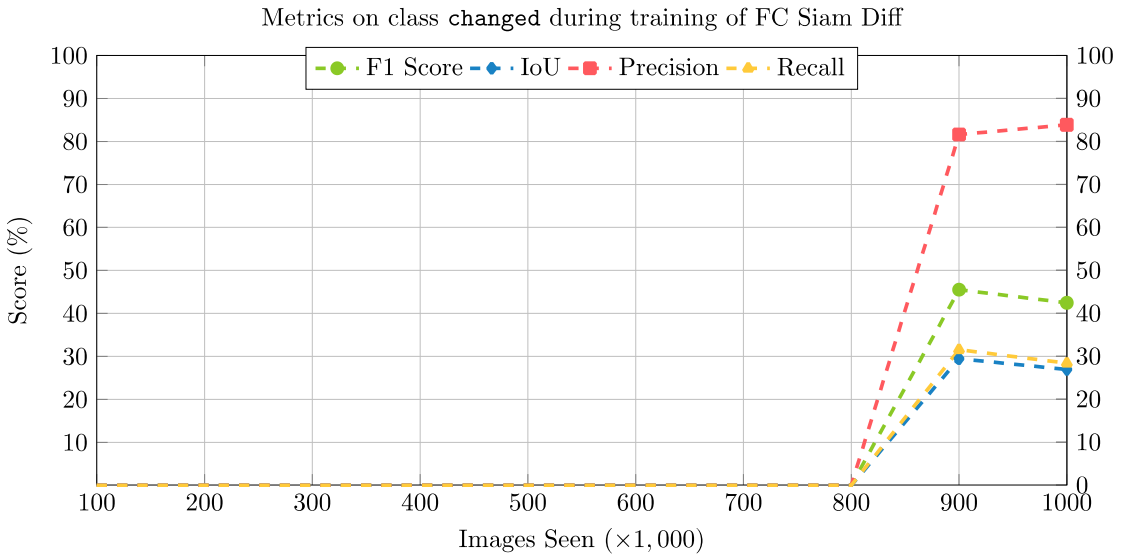
Model Training - Statistics

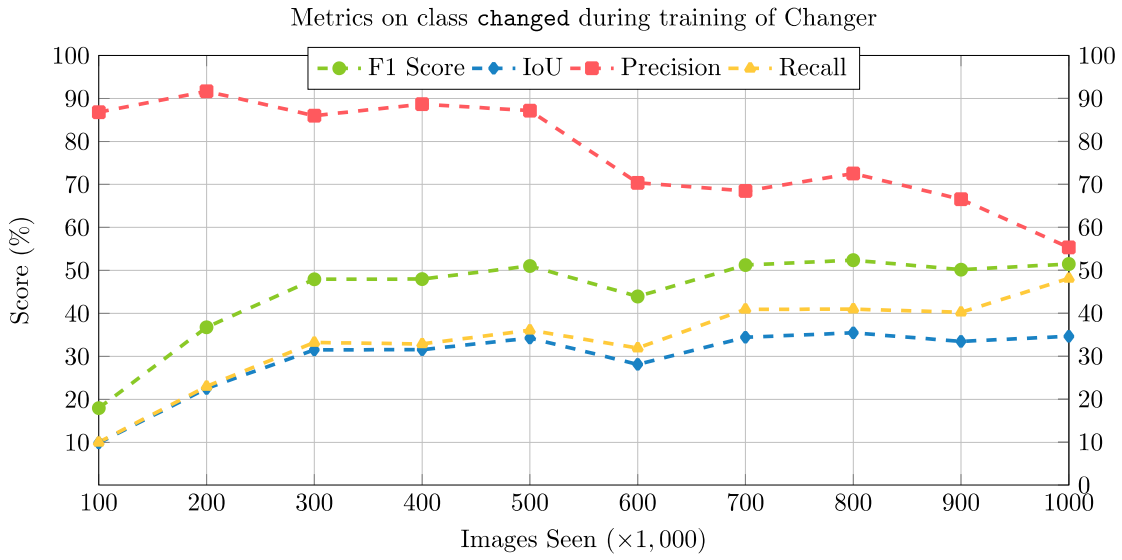
Model	**Batch Size**	**Max Iterations**	**Training Time** (hh:mm)	**Time** to best checkpoint (hh:mm)	**Iterations** to best checkpoint (×1 000)	**Pre-trained?**
FC-EF	32	31 250	03:40	00:22	100	✗
FC-Siam-Conc	24	41 667	05:01	05:01	1 000	✗
FC-Siam-Diff	24	41 667	05:00	04:30	900	✗
ChangeStar	24	41 667	08:12	08:12	1 000	✓
ChangeFormer	24	41 667	06:03	06:03	1 000	✓
Changer	32	31 250	05:13	04:10	800	✗
TTP	2	500 000	91:33	59:30	650	✓

Model Training – Non-Pretrained





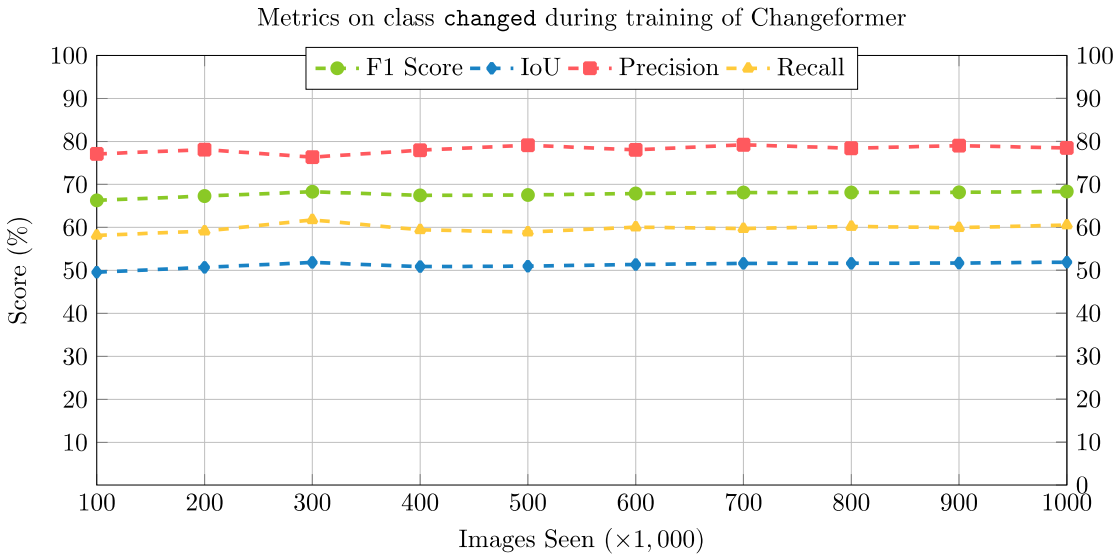


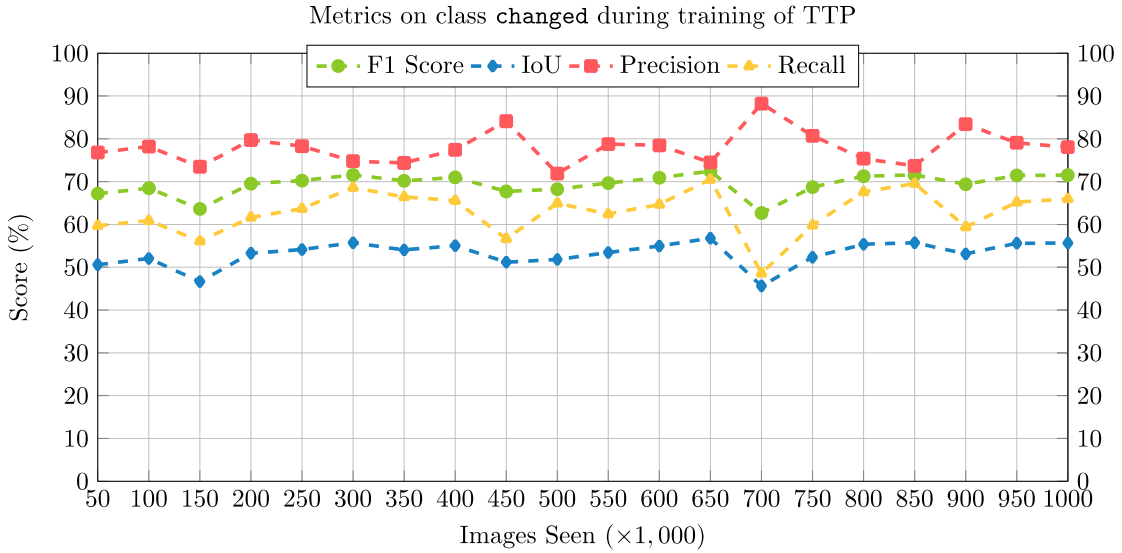


Recall: How well model identifies changed regions

Precision: How well model ignores irrelevant changes

Model Training – Pretrained

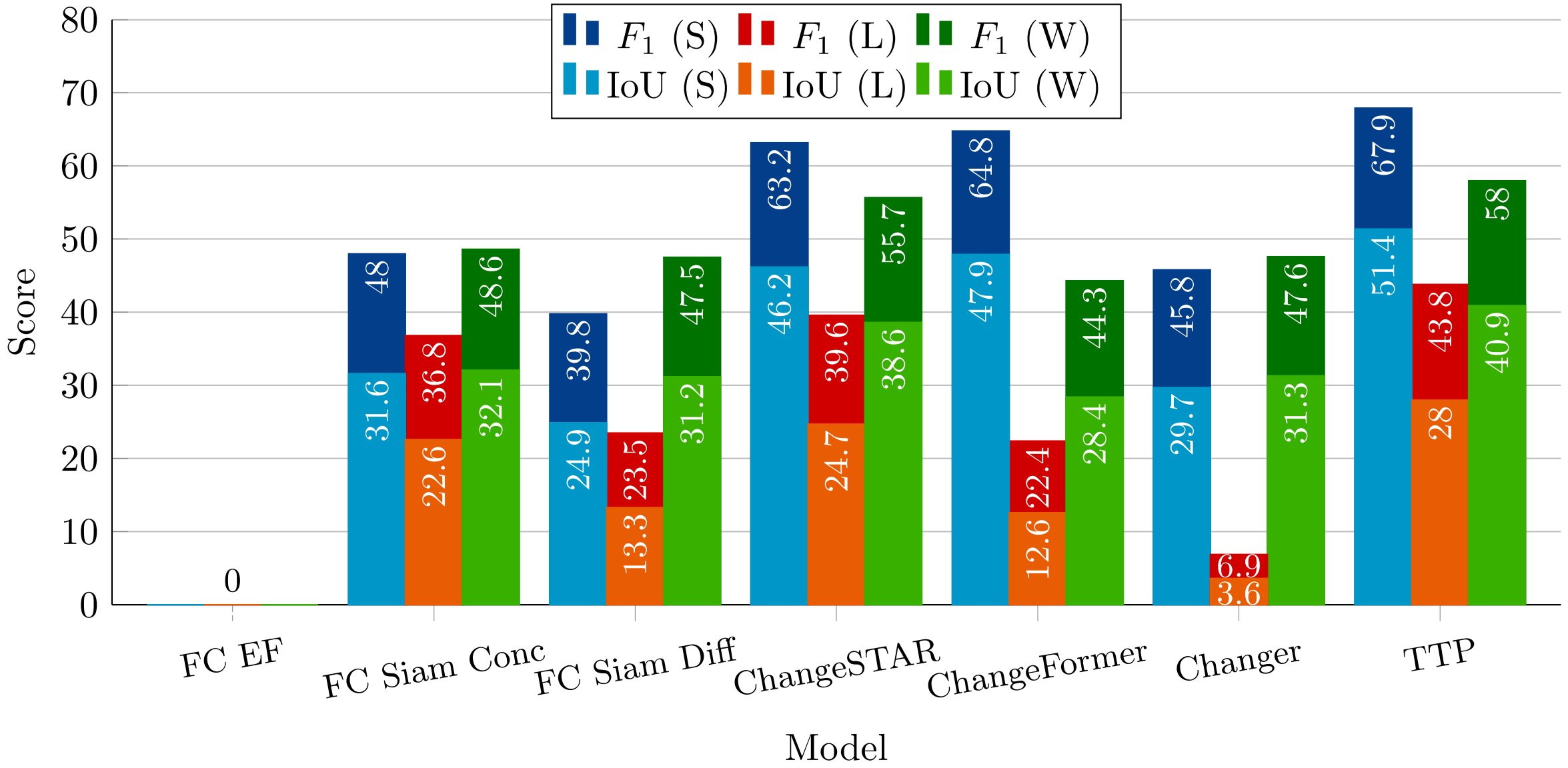




Recall: How well model identifies changed regions

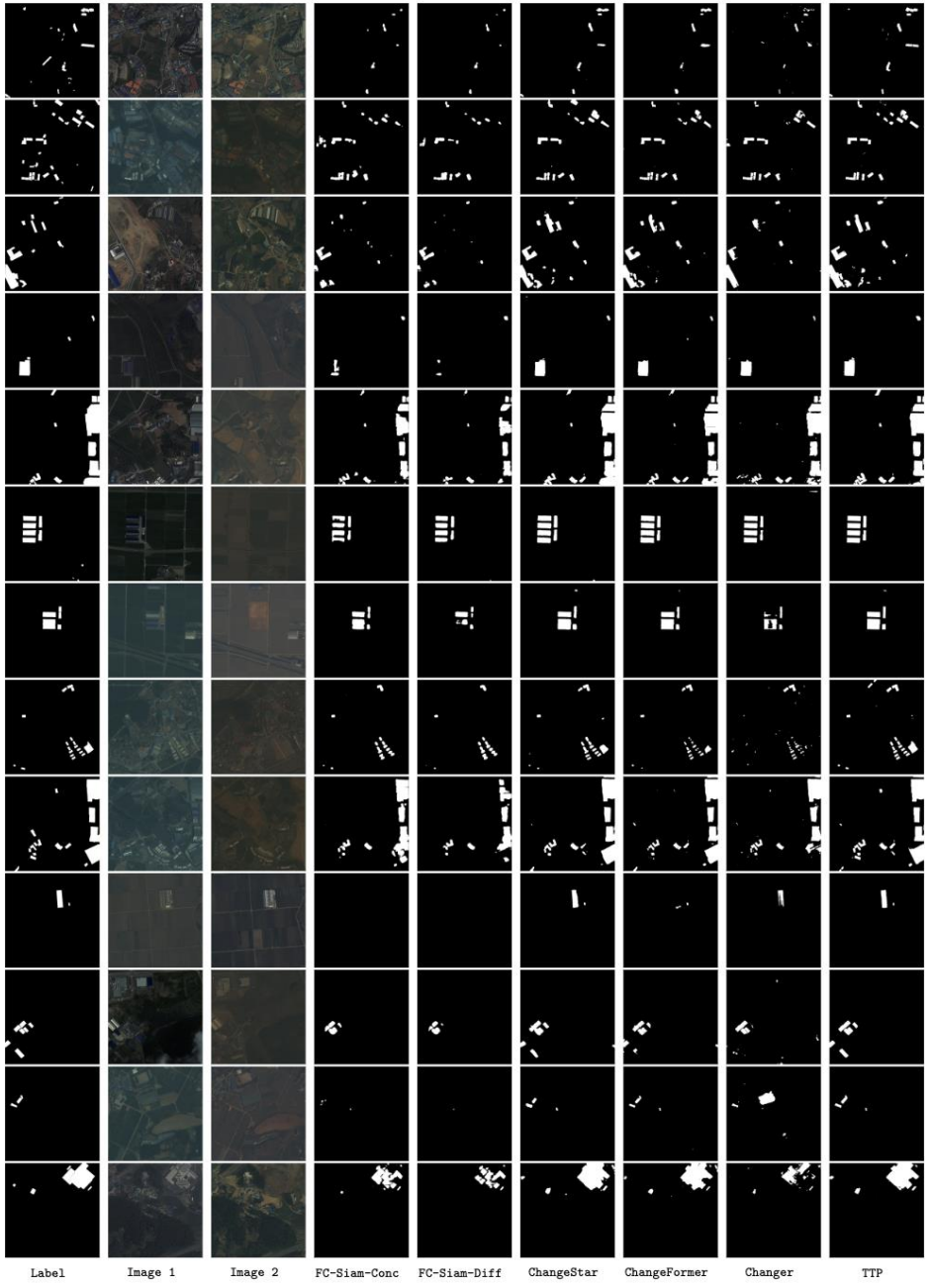
Precision: How well model ignores irrelevant changes

Quantitative Evaluation



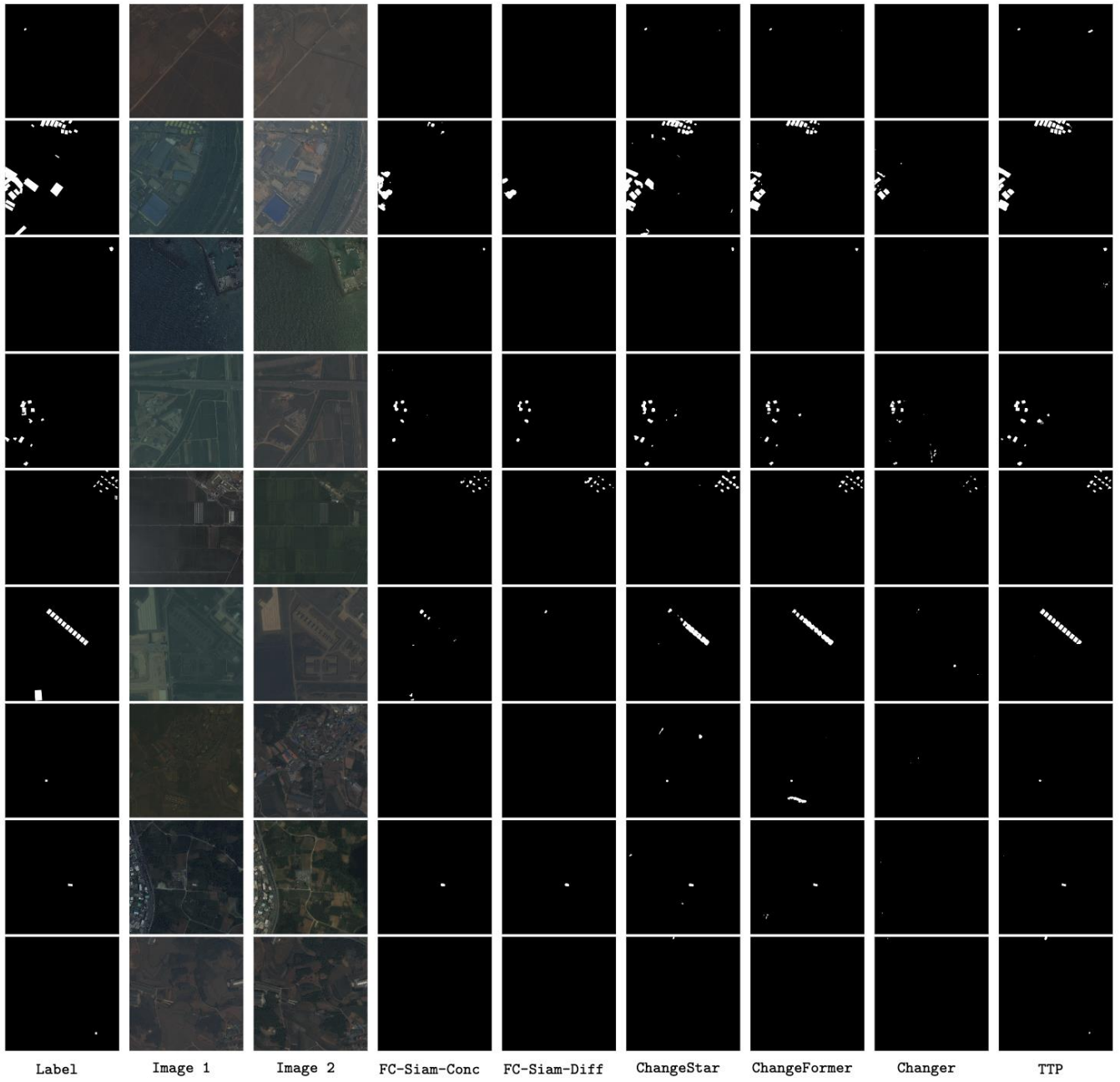
Qualitative Evaluation | S2Looking

Overall strong performance



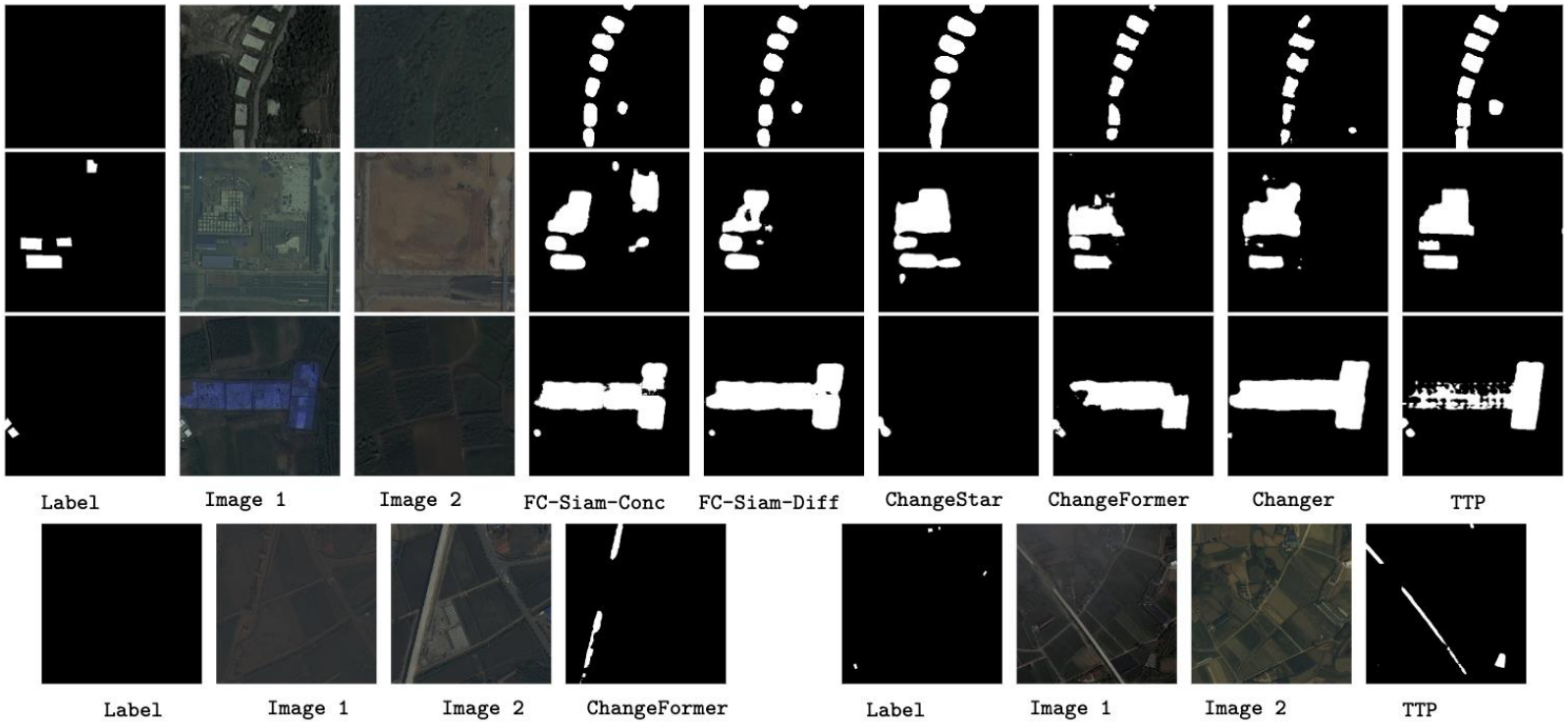
Qualitative Evaluation | S2Looking

Also *small changes* are identified well



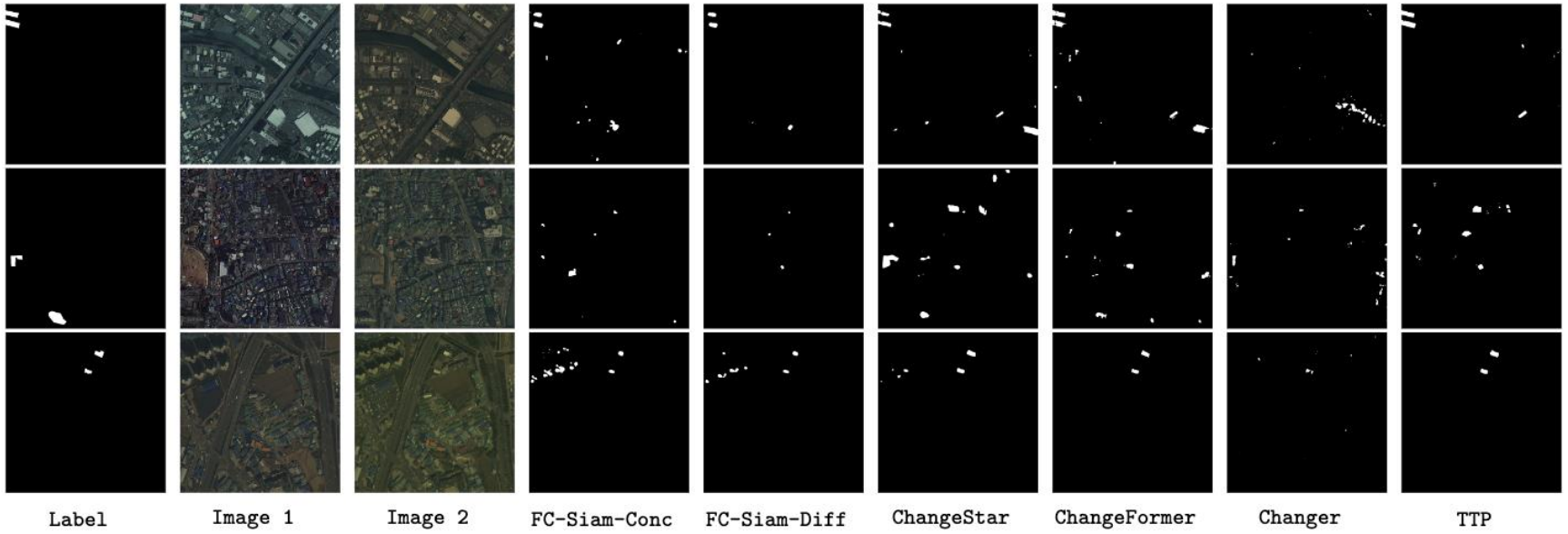
Qualitative Evaluation | S2Looking

False positives: construction sites, streets



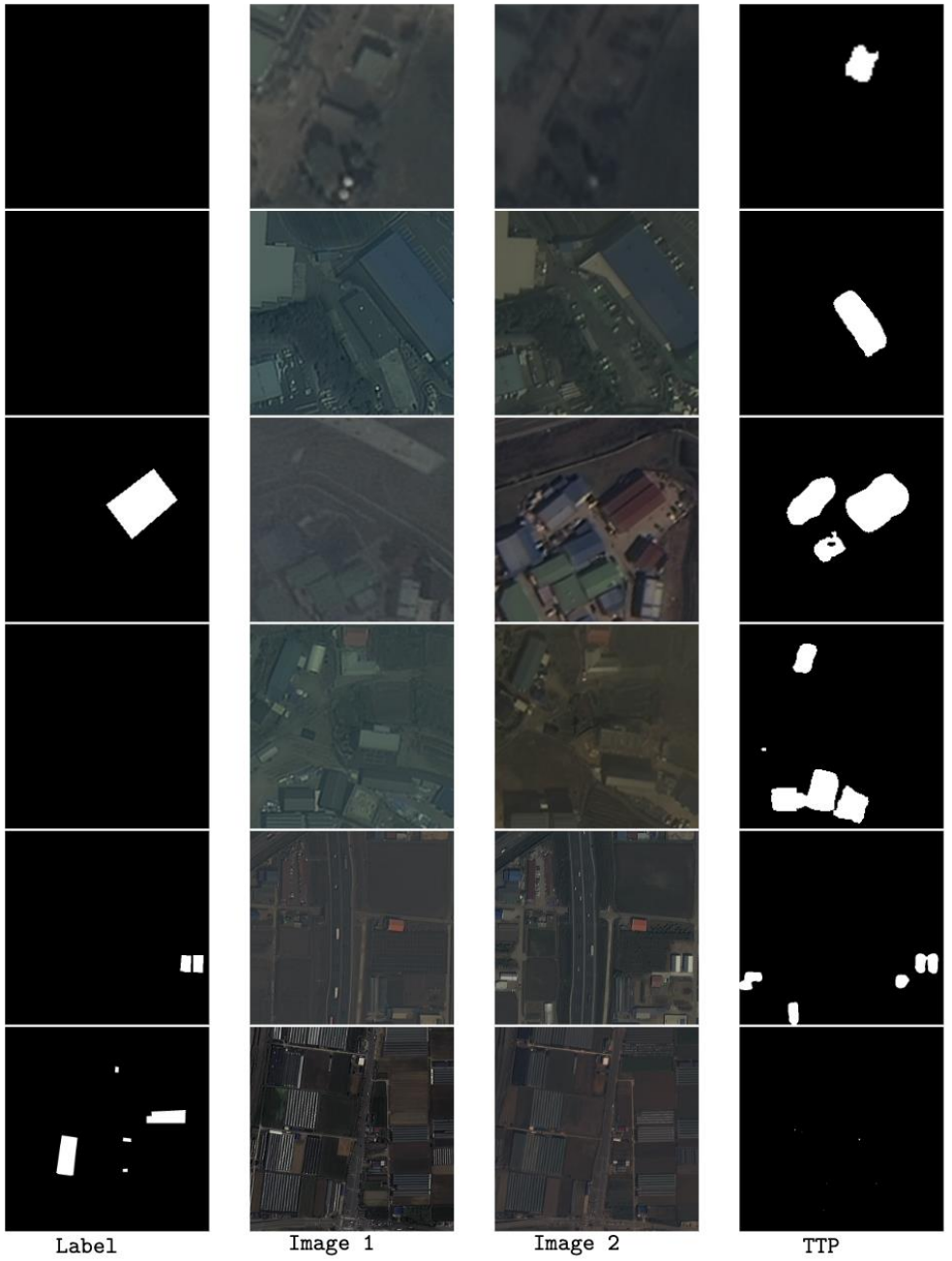
Qualitative Evaluation | S2Looking

Large off-nadir angles challenging



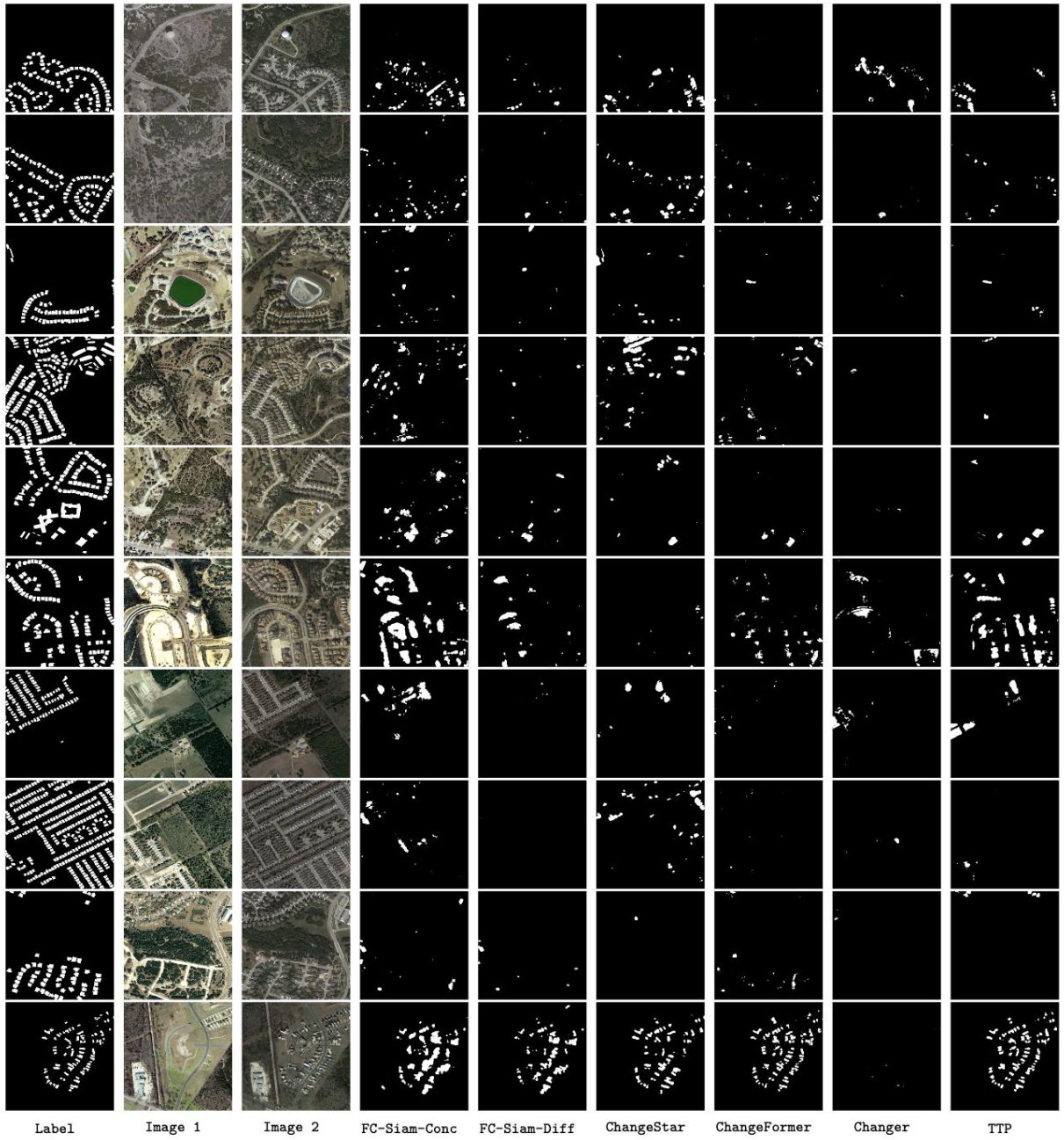
Qualitative Evaluation | S2Looking

Inconsistent or incorrect labels
→ TTP better



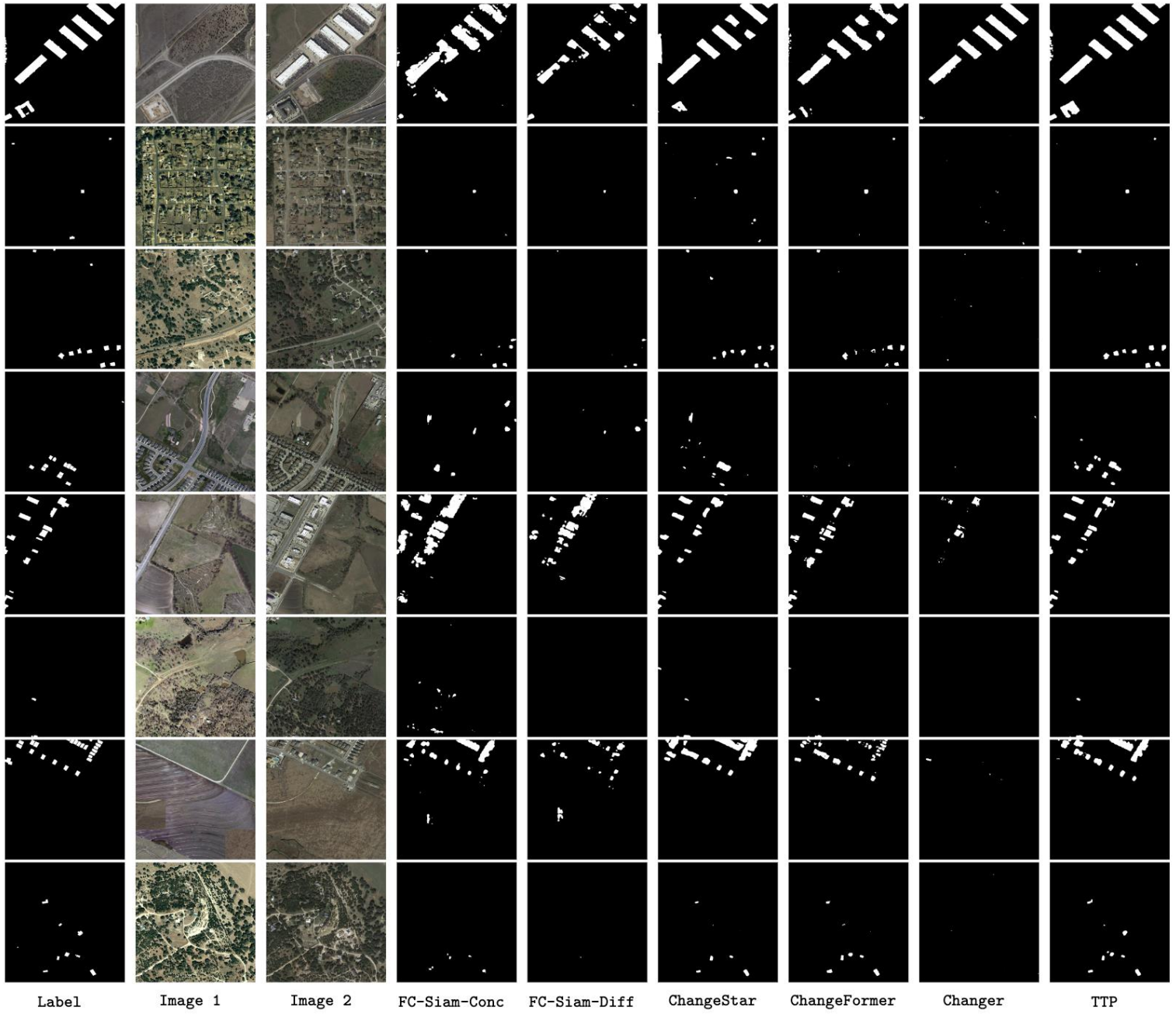
Qualitative Evaluation | LEVIR-CD

More changes
→ Difficult

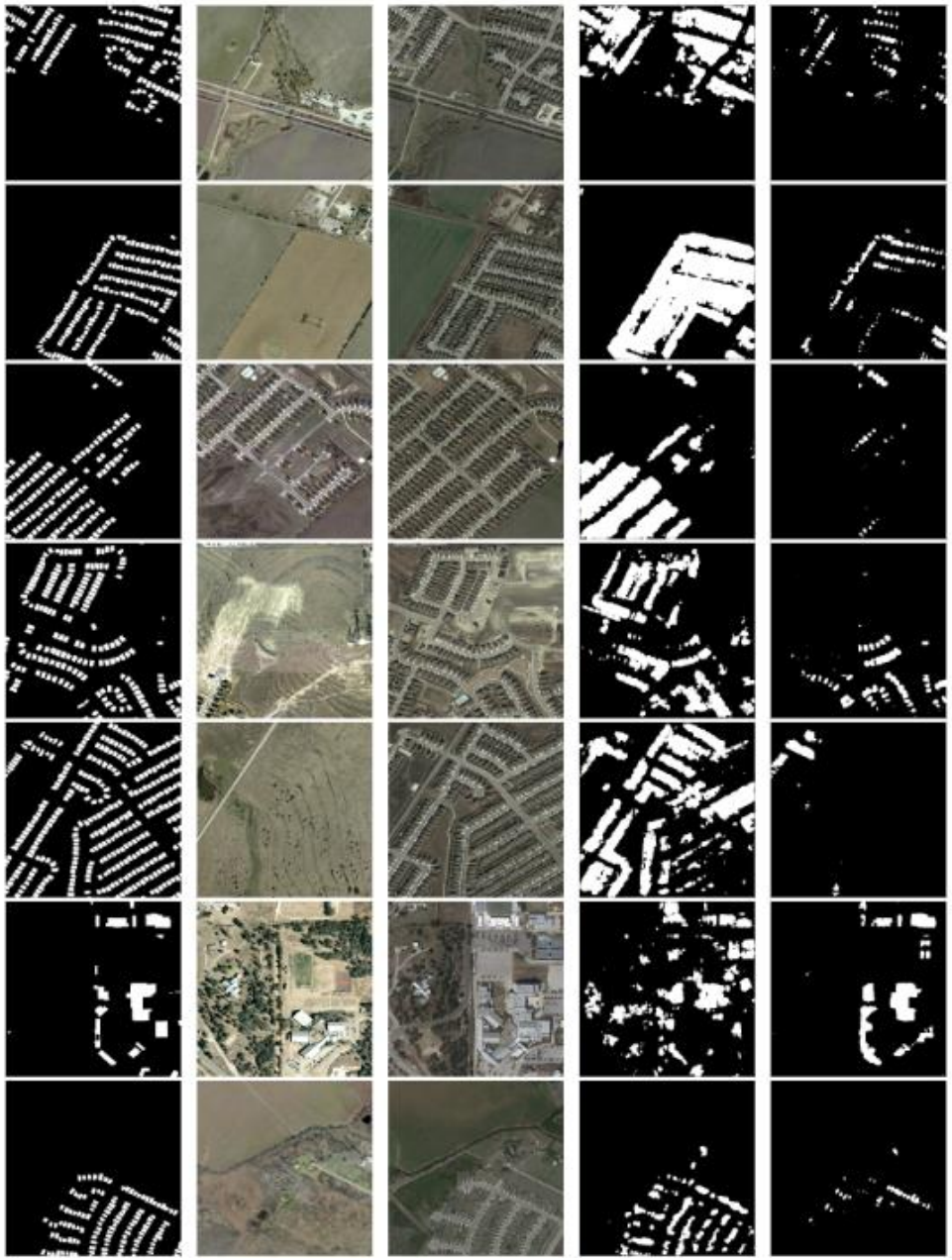


Qualitative Evaluation | LEVIR-CD

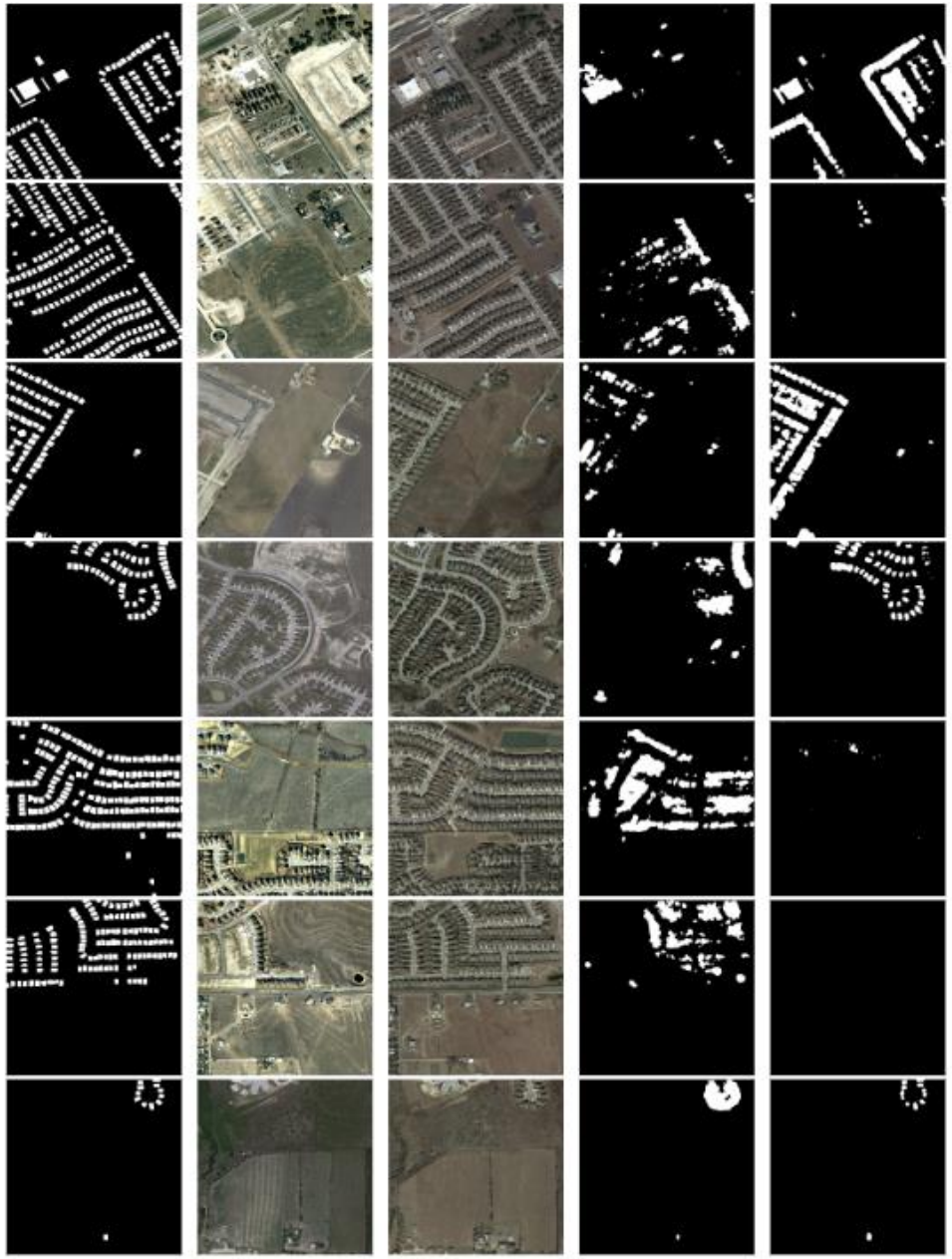
Less changes → Easier



Qualitative Evaluation | LEVIR-CD

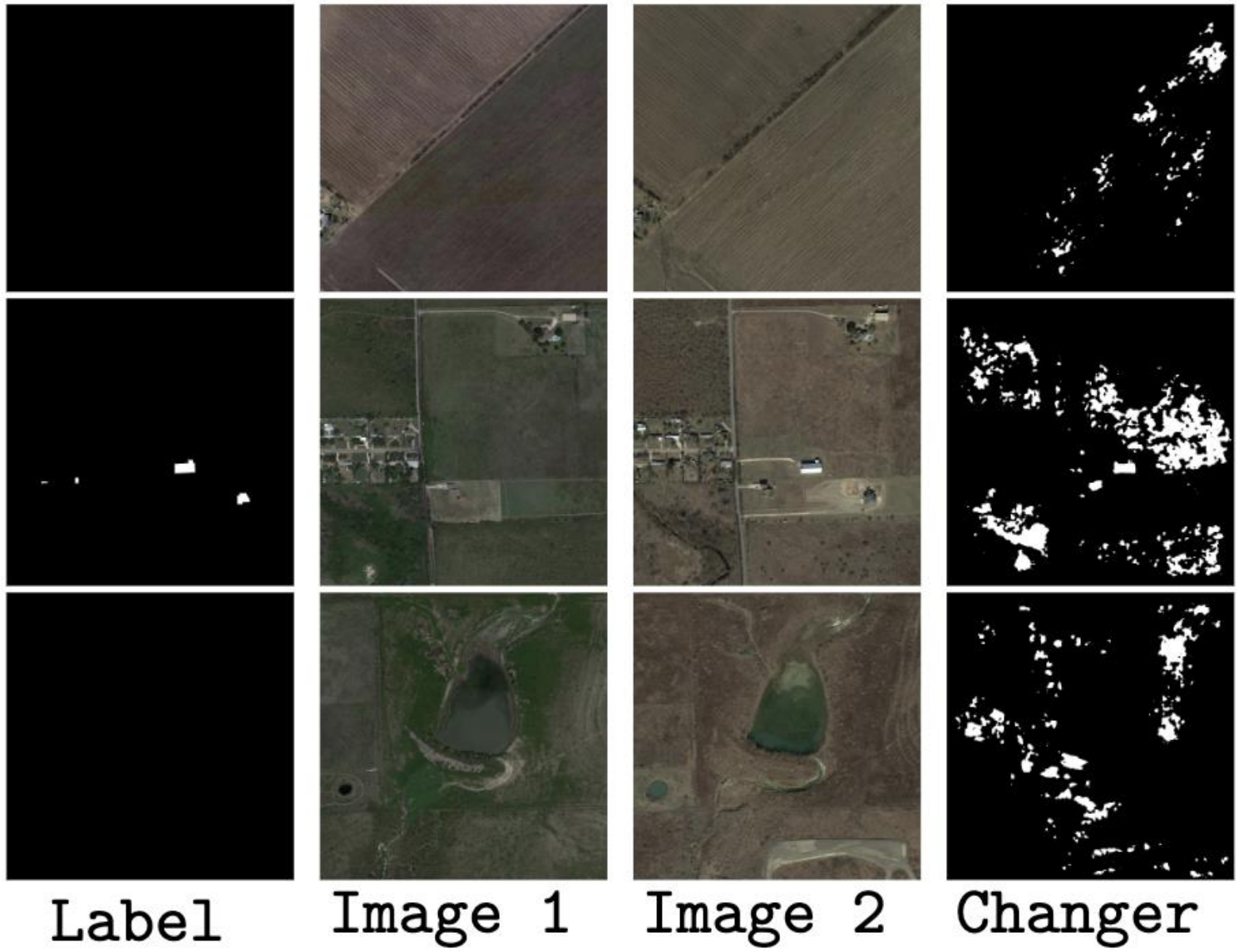


Siam-
Conc
vs.
TTP



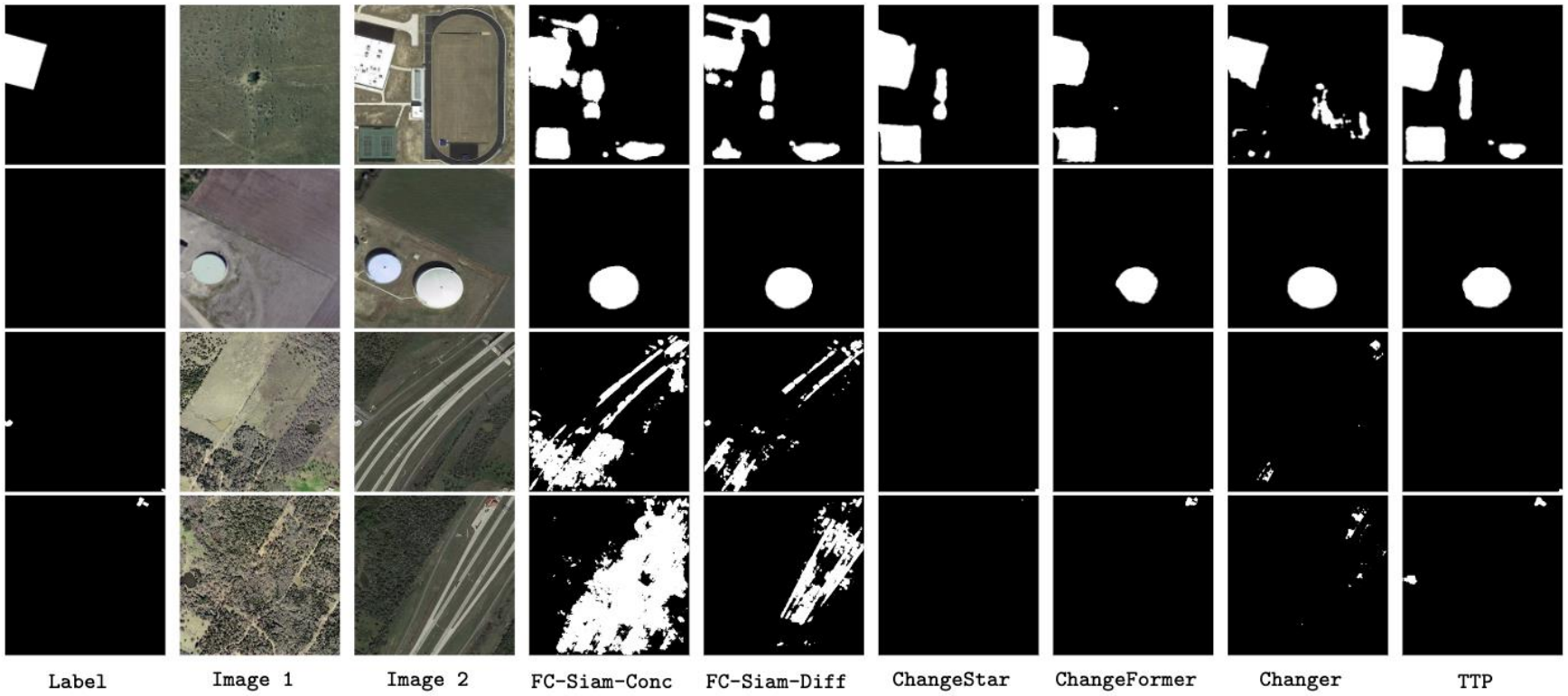
Qualitative Evaluation | LEVIR-CD

Changer hallucinates



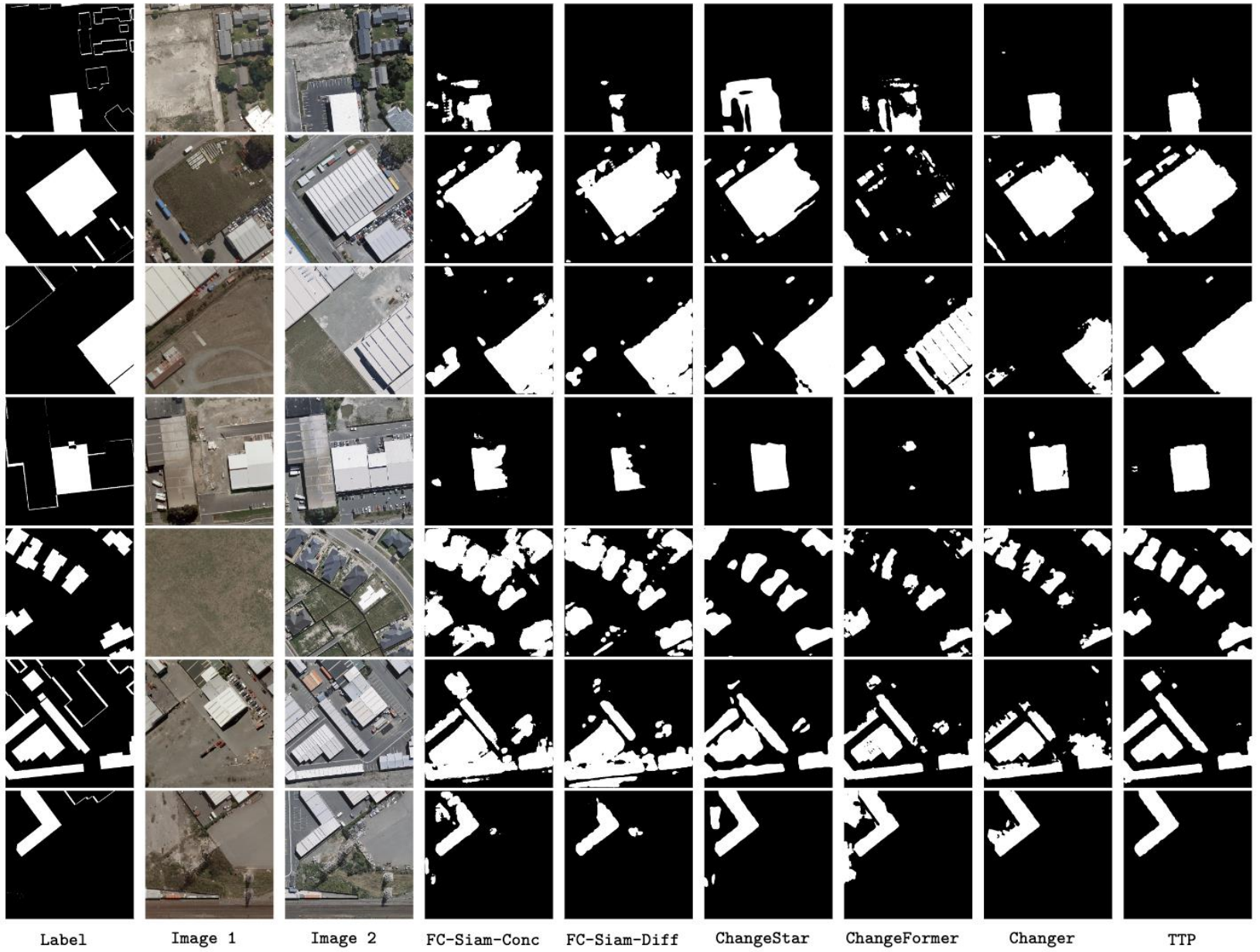
Qualitative Evaluation | LEVIR-CD

False positives: Tennis court, silo, street



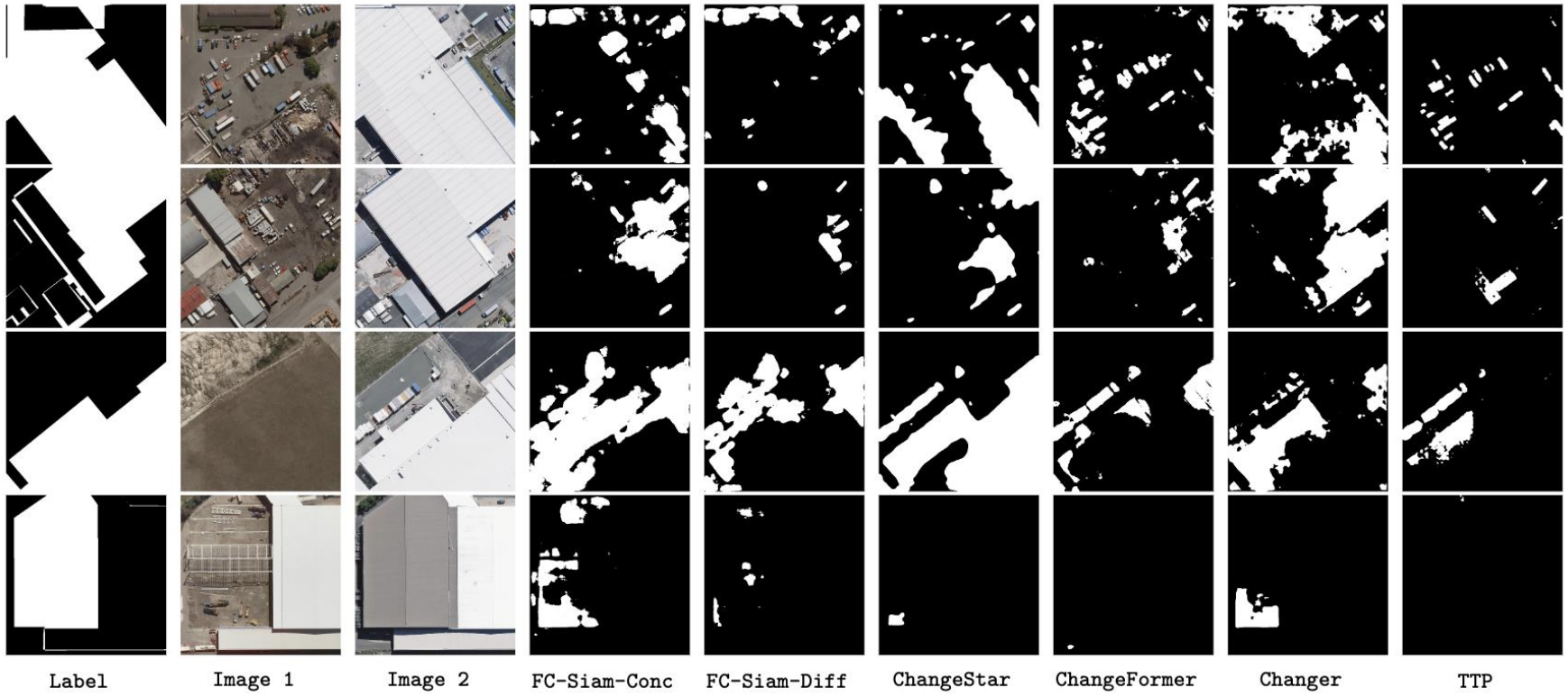
Qualitative Evaluation | WHU-CD

Generally well, TTP slightly better



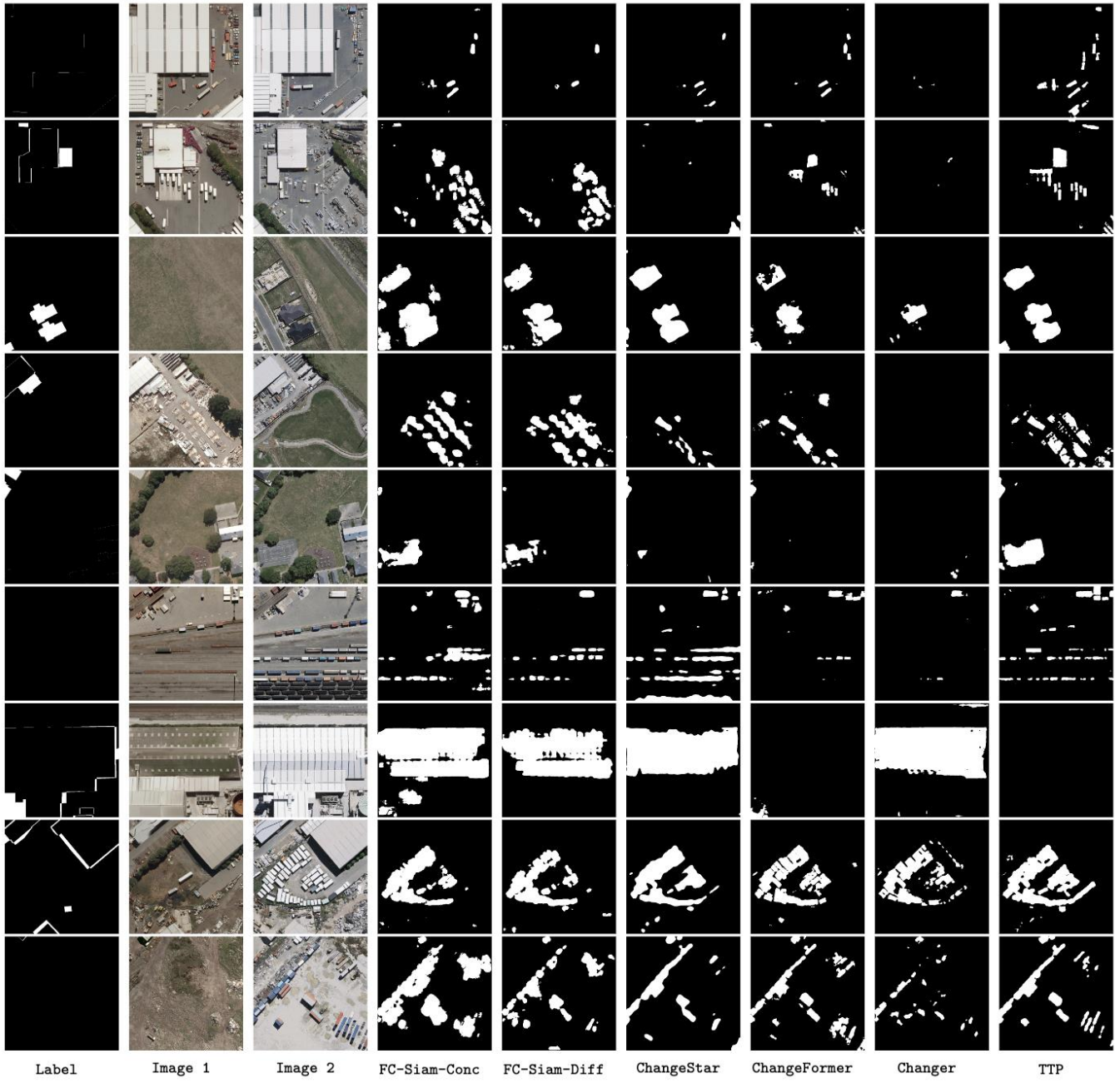
Qualitative Evaluation | WHU-CD

Models struggle with huge changes



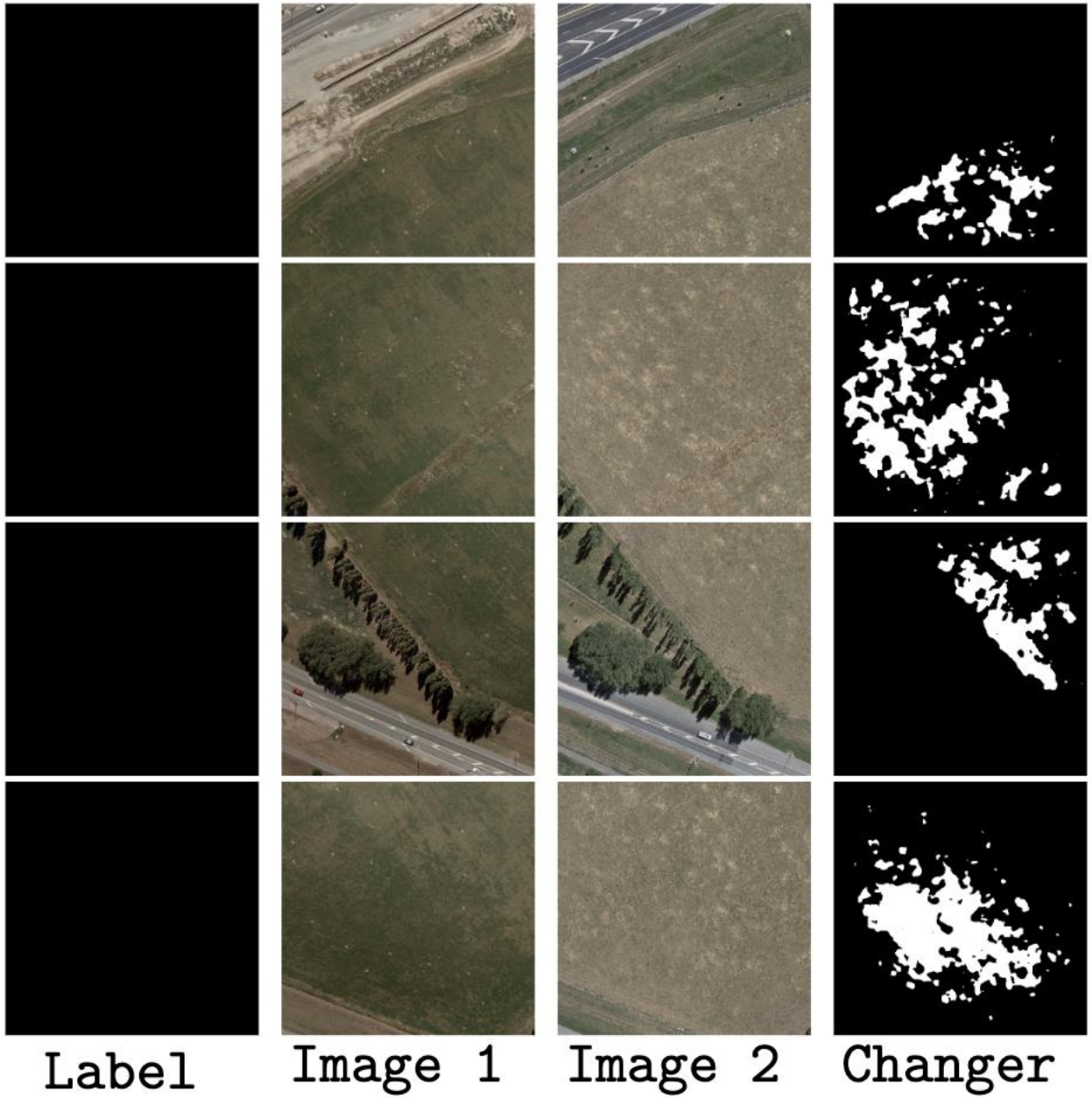
Qualitative Evaluation | WHU-CD

False positives:
Containers,
trucks, wagons



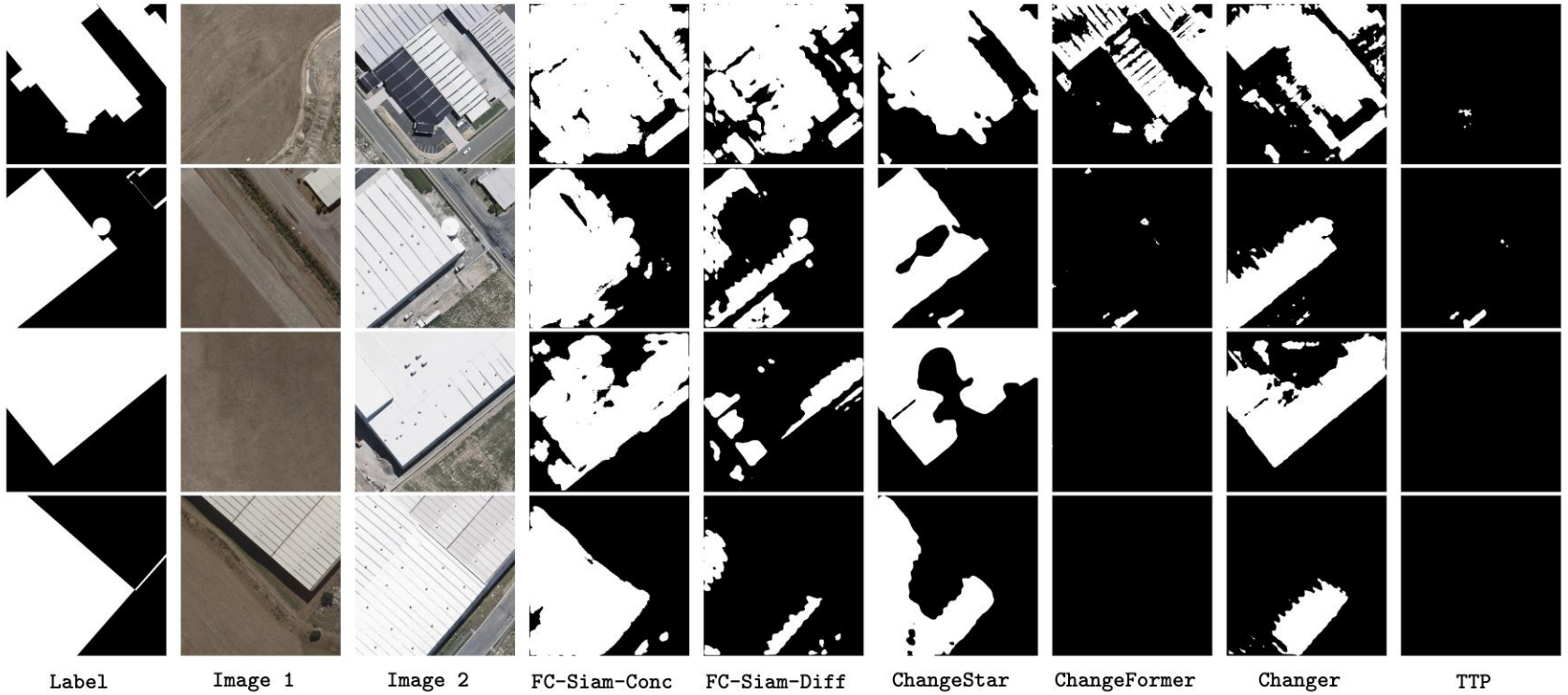
Qualitative Evaluation | WHU-CD

Changer hallucinates



Qualitative Evaluation | WHU-CD

TTP especially struggles with huge changes



Conclusion

CD is a highly relevant task

Some models were explained and tested

Training on S2Looking is not enough to obtain models that perform well across all datasets