False False-Positives

During the evaluation, we noticed that many of the samples categorized as False-Positives (FP) were actual traffic signs. Therefore, we also reported the metrics disregarding these *False False-Positives* (FFP) by manually going through all the False-Positives and verifying which ones were actual traffic signs. Table I shows a summary of how many images in each of the 10 runs were classified as *False False-Positives* and the impact on the metrics. After that, we show all the images that were disregarded.

TABLE I
IMPACT OF FALSE FALSE-POSITIVES ON THE METRICS

Runs	# FP	# FFP	Precision (±Δ)	F1-score (±Δ)
#1 (Figure 1)	18	6	**96.40%** (+1.71%)	0.9251 (+0.0079)
#2 (Figure 2)	38	19	94.56% (+4.88%)	0.9296 (+0.0242)
#3 (Figure 3)	39	19	94.44% (+4.73%)	0.9431 (+0.0242)
#4 (Figure 4)	32	15	95.07% (+3.96%)	0.9292 (+0.0193)
#5 (Figure 5)	27	13	95.92% (+3.50%)	0.9347 (+0.0169)
#6 (Figure 6)	32	14	94.97% (+3.57%)	**0.9458** (+0.0181)
#7 (Figure 7)	40	19	94.10% (+4.77%)	0.9344 (+0.0241)
#8 (Figure 8)	39	15	93.26% (+3.77%)	0.9261 (+0.0190)
#9 (Figure 9)	37	17	94.33% (+4.33%)	0.9328 (+0.0217)
#10 (Figure 10)	38	15	93.54% (+3.78%)	0.9289 (+0.0190)
Average	–	–	94.66% (+3.90%)	0.9330 (+0.0195)

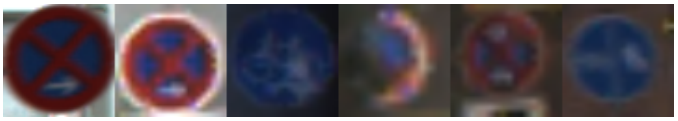

Fig. 1. Proposed Method run #1

Fig. 2. Proposed Method run #2

Fig. 3. Proposed Method run #3

Fig. 4. Proposed Method run #4

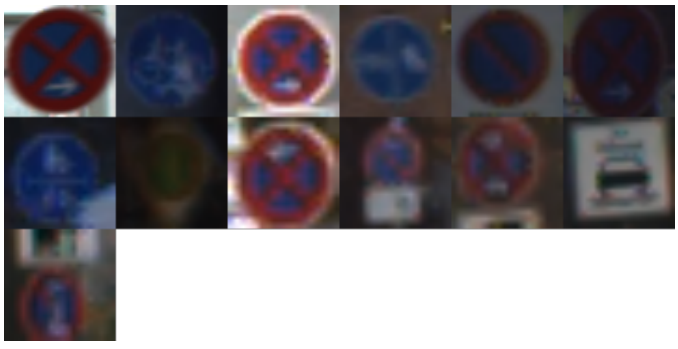

Fig. 5. Proposed Method run #5

Fig. 6. Proposed Method run #6

Fig. 7. Proposed Method run #7

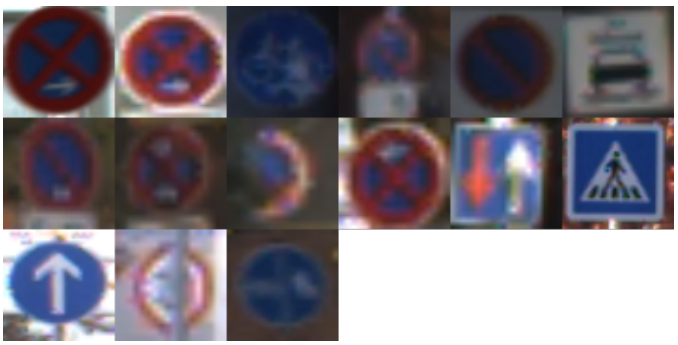

Fig. 8. Proposed Method run #8

Fig. 9. Proposed Method run #9

Fig. 10. Proposed Method run #10