# Guided Retraining: ablation study

In this document, we provide the results of our ablation study about the importance of Guided Retraining's components. Specifically, we have considered two settings: (1)- Instead of training four Models on combinations of TP, FP, TN and FN, we studied the impact of considering only two Models (i.e., one is trained on TP and FP and the other is trained on TN and FN). (2)- We kept the four Models, but we removed Model5 and trained the auxiliary classifier directly on the concatenation of the output of the four Models.

We report the results of these experiments below in the Table:

| | **Only 2 sub-Models** | | | **Without Model5** | | | **GuidedRetraining** | | |
|---|---|---|---|---|---|---|---|---|---|
| | F1 (%) | ΔErr | Errors reduction | F1 (%) | ΔErr | Errors reduction | F1 (%) | ΔErr | Errors reduction |
| DREBIN | 90.10 | 169 | -31.30% | 93.20 | -53 | 9.81% | 93.44 | -69 | **12.78%** |
| Reveal | 90.19 | -189 | 27.96% | 85.38 | 14 | -2.07% | 91.44 | -248 | **36.69%** |
| MaMaF | 96.34 | -233 | 31.07% | 96.14 | -208 | 27.73% | 96.55 | -264 | **35.20%** |
| MaMaP | 96.64 | 65 | -14.64% | 96.85 | 31 | -6.98% | 97.13 | -9 | **2.03%** |
| MalscanA | 91.42 | -223 | 37.29% | 91.11 | -214 | 35.79% | 91.83 | -243 | **40.64%** |
| MalscanCO | 91.70 | -177 | **32.90%** | 88.35 | -57 | 10.59% | 91.66 | -177 | **32.90%** |