

RESULTS OF EACH MODEL (VOLO AND SWIN) TRAINED ON DIFFERENT DATASETS

FTTQ (FAKE & REAL) DATASET

	VOLO	SWIN
AUC	0.853	0.811
EER	0.2434	0.2857
TDR_01	0.0122	0.1365
TDR_001	0.0122	0.1365
PBCA	0.7607	0.7183

DFFD DATASET

	VOLO	SWIN
AUC	0.9808	0.9833
EER	0.0509	0.0477
TDR_01	0.9197	0.9260
TDR_001	0.9291	0.7999
PBCA	0.9649	0.9648

SELF CREATED DATASET

	VOLO	SWIN
AUC	0.998	0.967
EER	0.0214	0.068
TDR_01	0.8914	0.0318
TDR_001	0.6935	0.0856
PBCA	0.9793	0.9390

STARGANN DATASET

	VOLO	SWIN
AUC	1	0.5
EER	0	0
TDR_01	1	0
TDR_001	1	0
PBCA	1	0

LLC_FSD DATASET

	VOLO	SWIN (0.3)	SWIN (0.5)
AUC	0.9791	0.5025	0.957
EER	0.0637	0.4968	0.0828
TDR_01	0.1413	0.00318	0.142
TDR_001	0.1413	0.0318	0.142
PBCA	0.9361	0.50955	0.92308

Here are the results after evaluating the best epoch for each model. I trained both VOLO D5 and SWIN (tiny) on each dataset. On the Stargann dataset, both models showed significant overfitting, despite the data being of high quality. I carefully checked the training, testing, and validation sets and confirmed there were no common images between them.

For the LLC_FASD dataset, which is referenced in the “Leveraging Swin Transformer for Robust Deepfake Detection” paper, I trained and tested both models. VOLO D5 performed better than SWIN (tiny) when the SWIN model’s dropout was set to 0.3.

To improve SWIN’s results, I increased the dropout to 0.5 and retrained the model on the same dataset. With this adjustment, the SWIN model performed better than before. These findings highlight the importance of tuning hyperparameters like dropout for achieving optimal model performance.

The graphs below display the model's accuracy, validation loss, and validation accuracy on their respective datasets for each epoch. The title at the top of each graph indicates what is being shown.

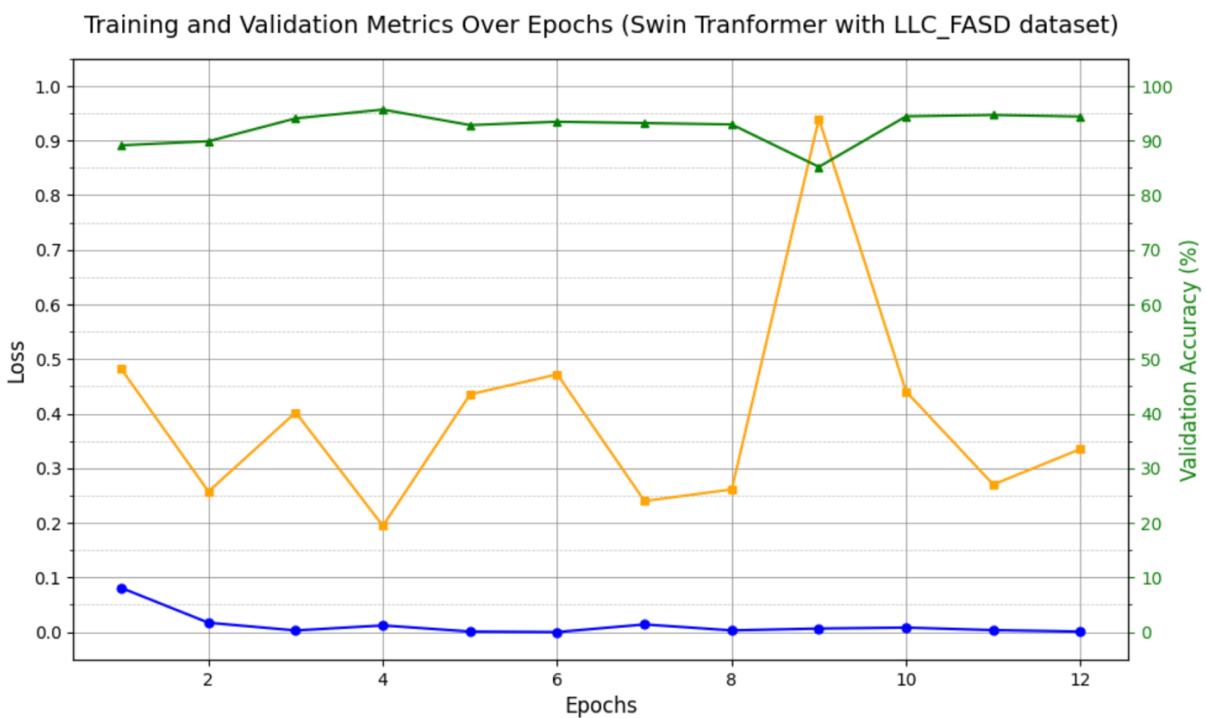
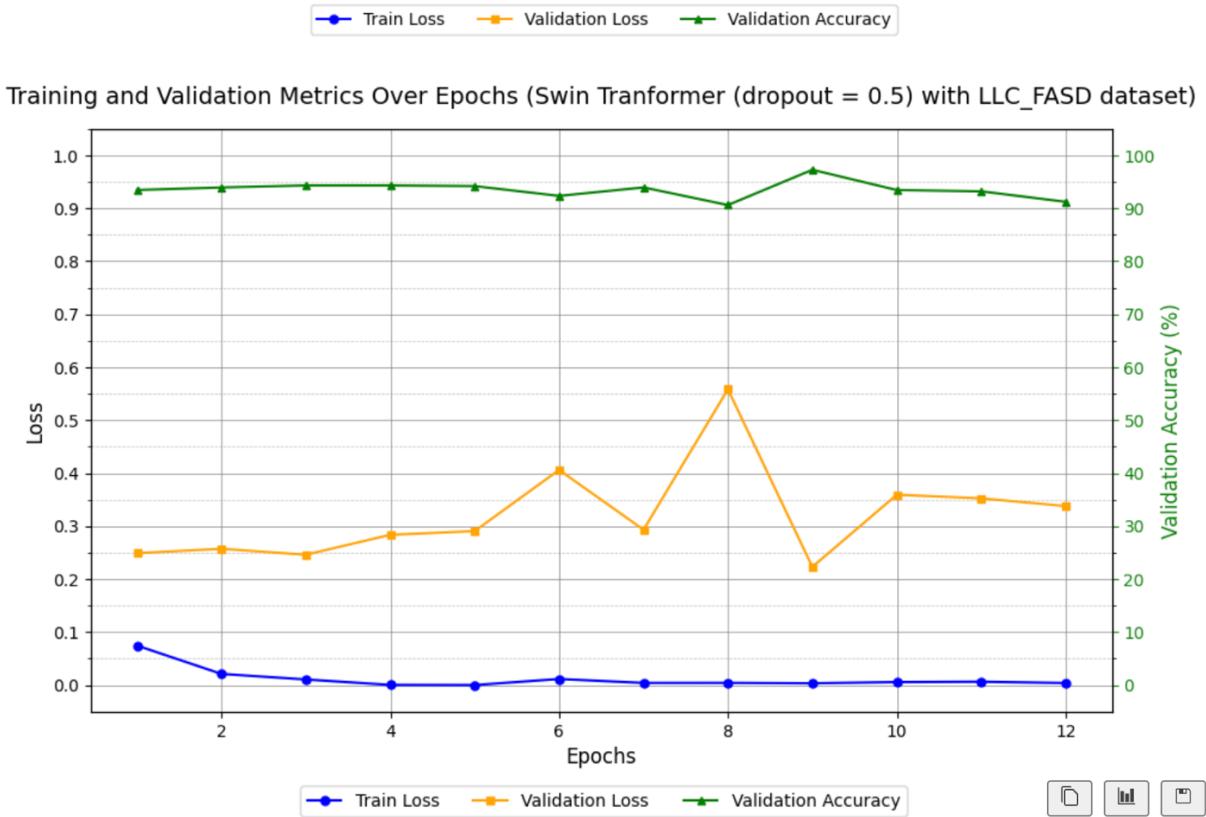
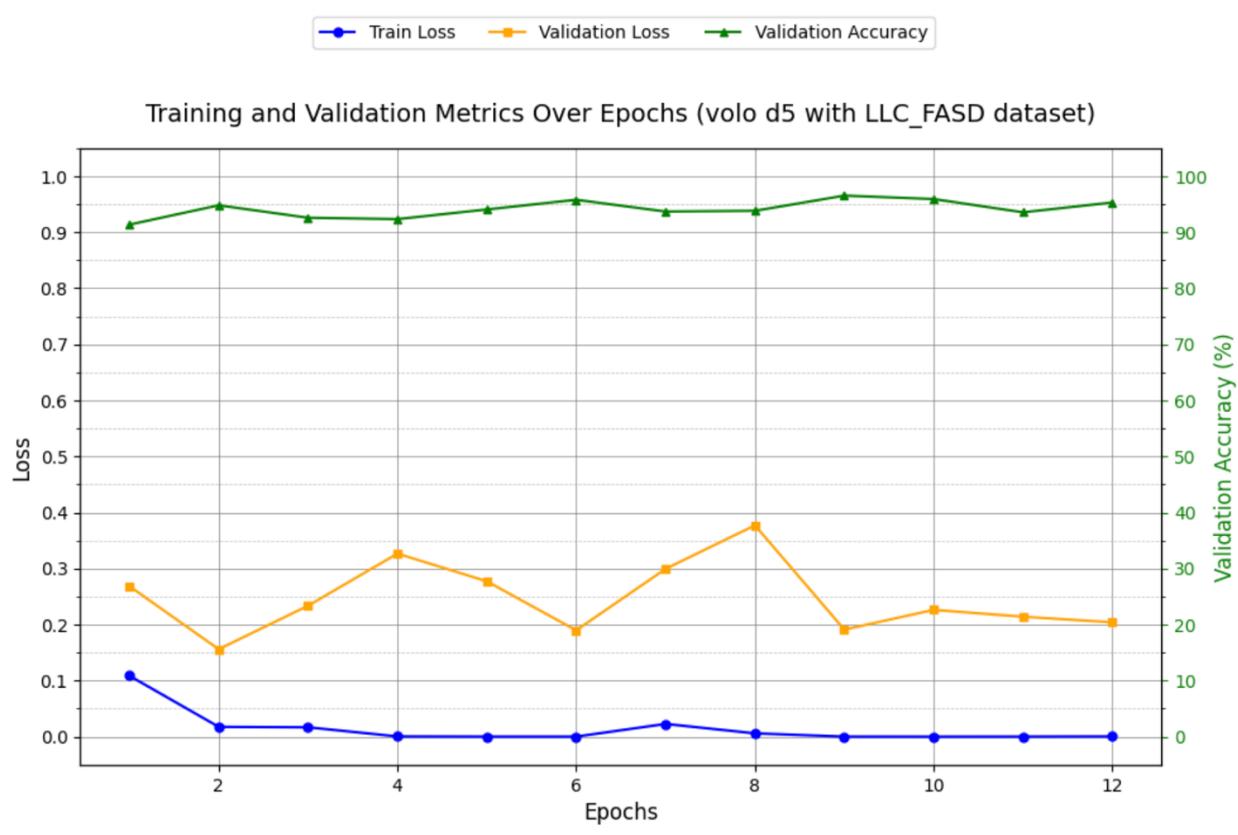
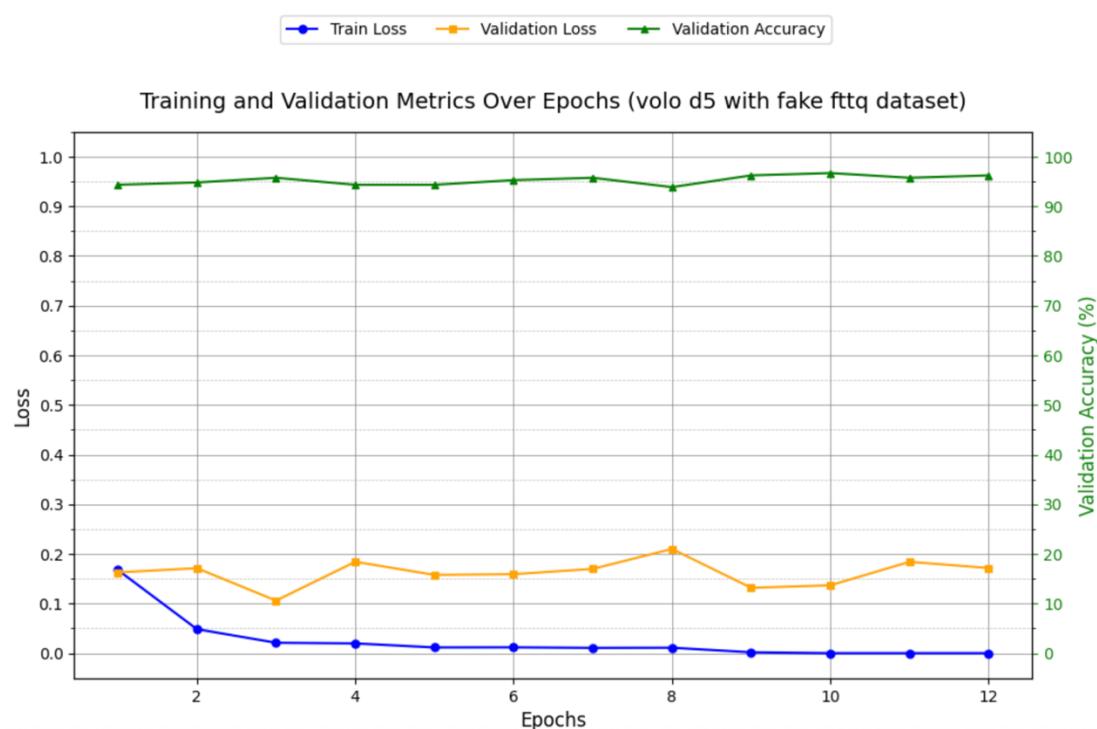
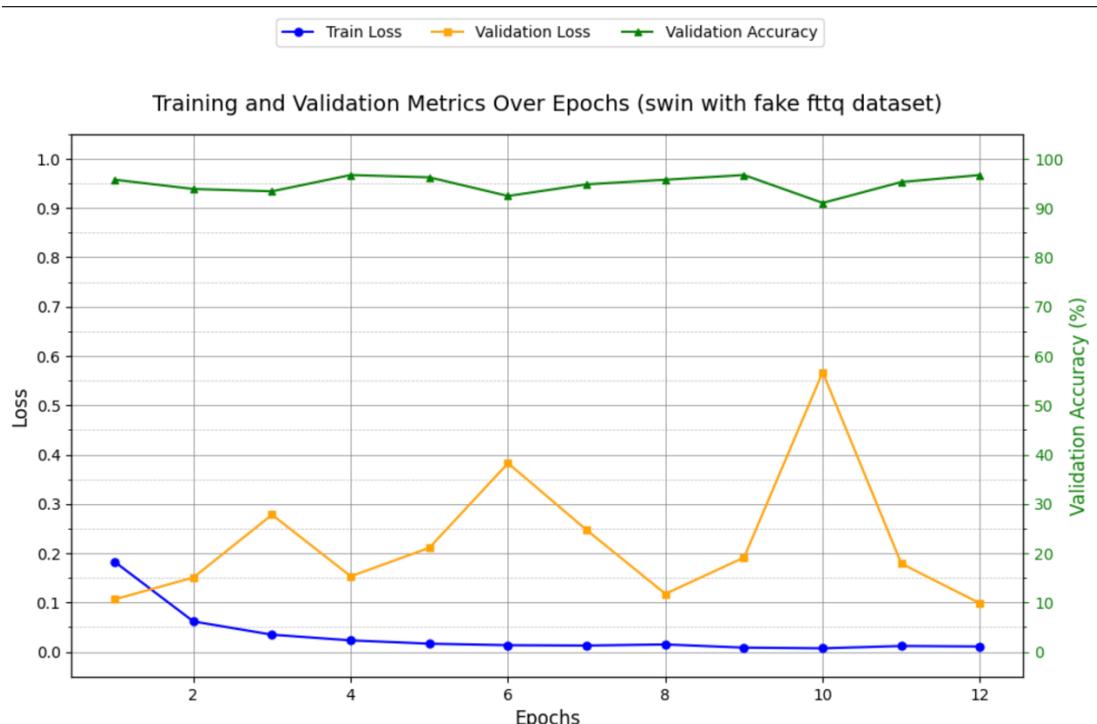


Figure: SWIN MODEL with dropout = 0.3

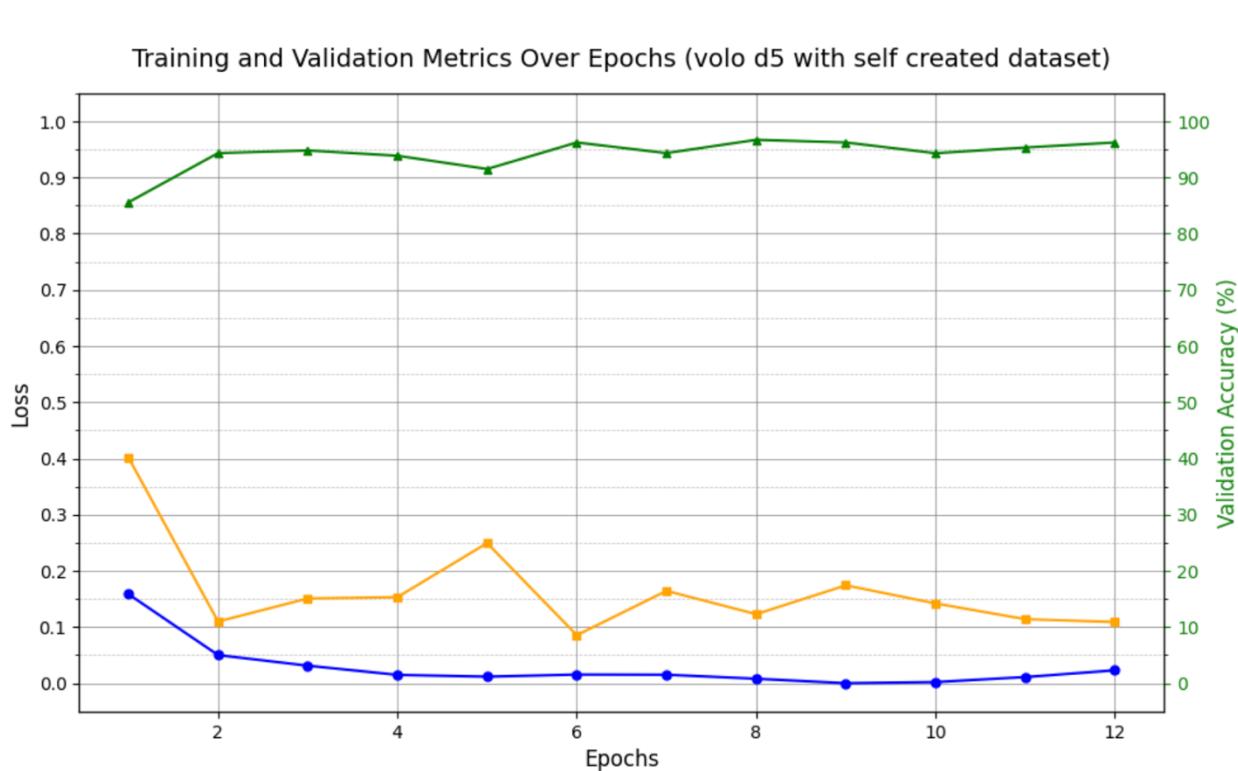
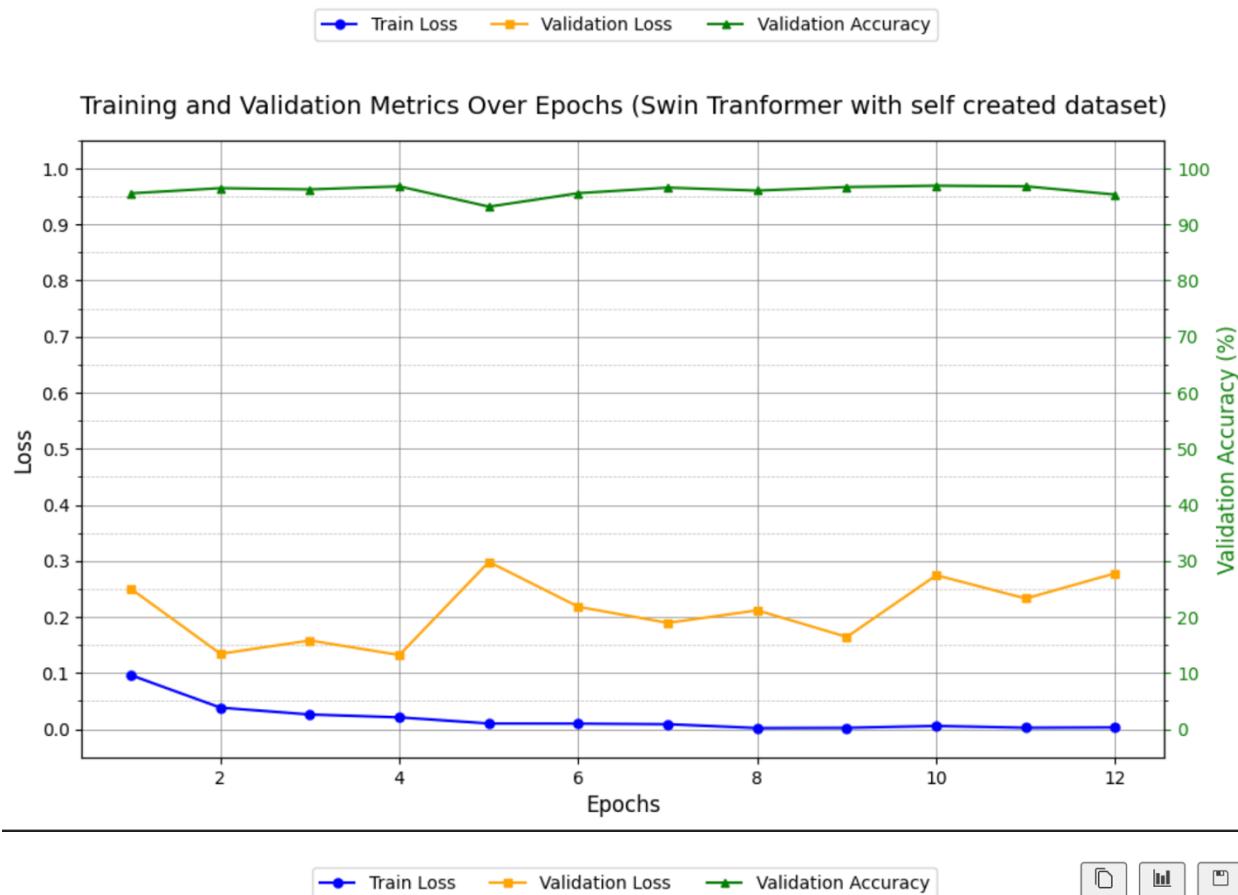
When training the Swin Model with a dropout rate of 0.3, the validation loss spiked sharply to 0.94 at the 9th epoch. However, after increasing the dropout rate to 0.5 and

retraining the model, there was no such drastic increase in validation loss compared to when dropout was set to 0.3.

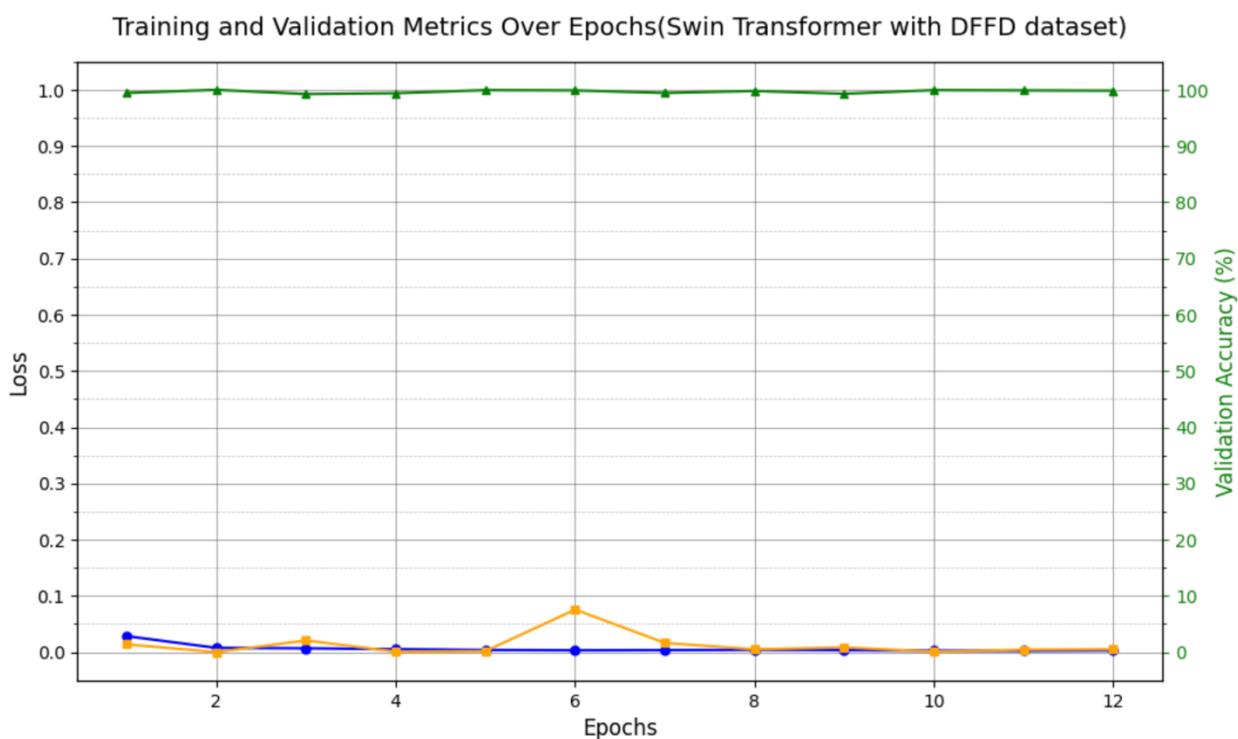
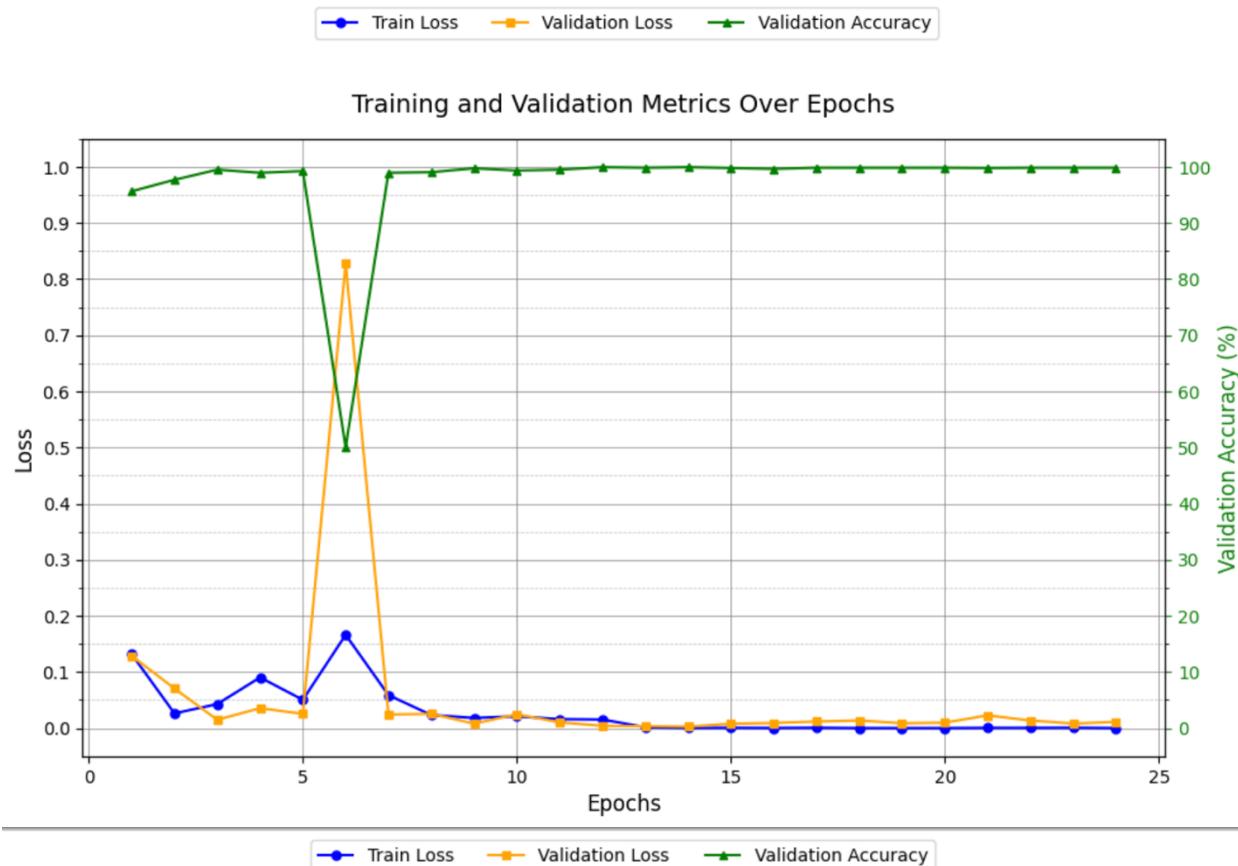




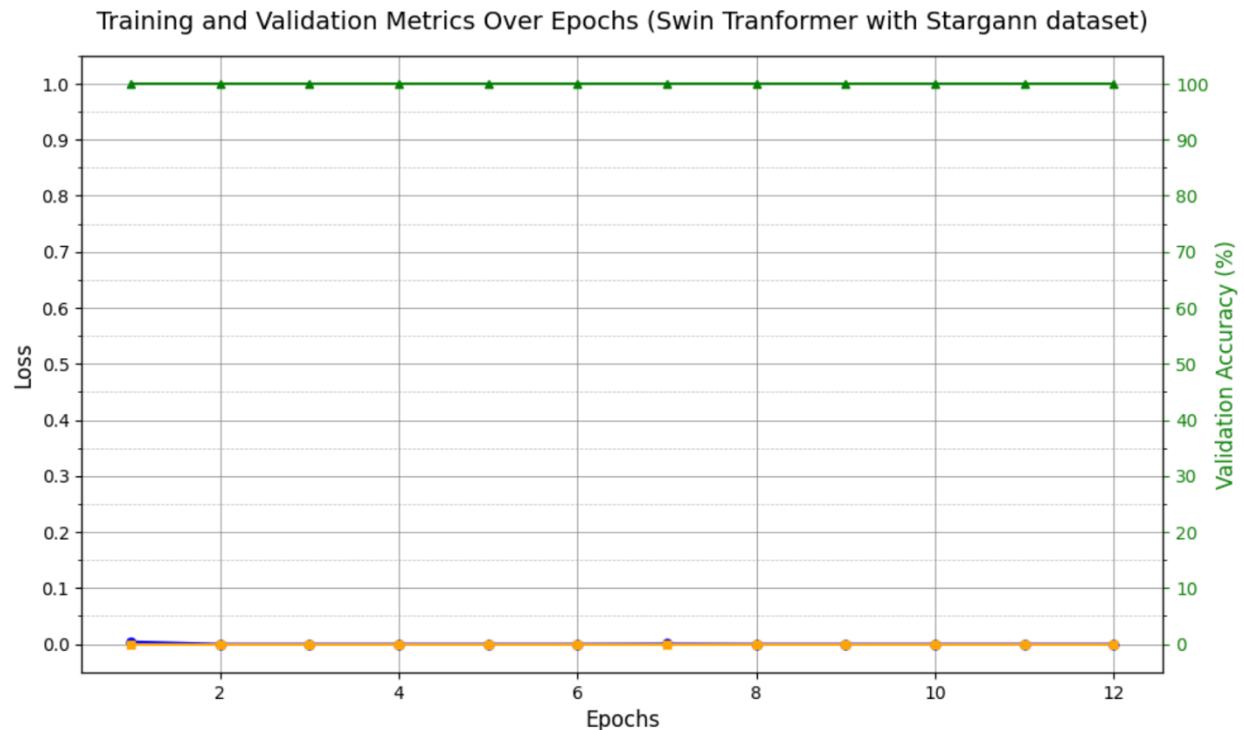
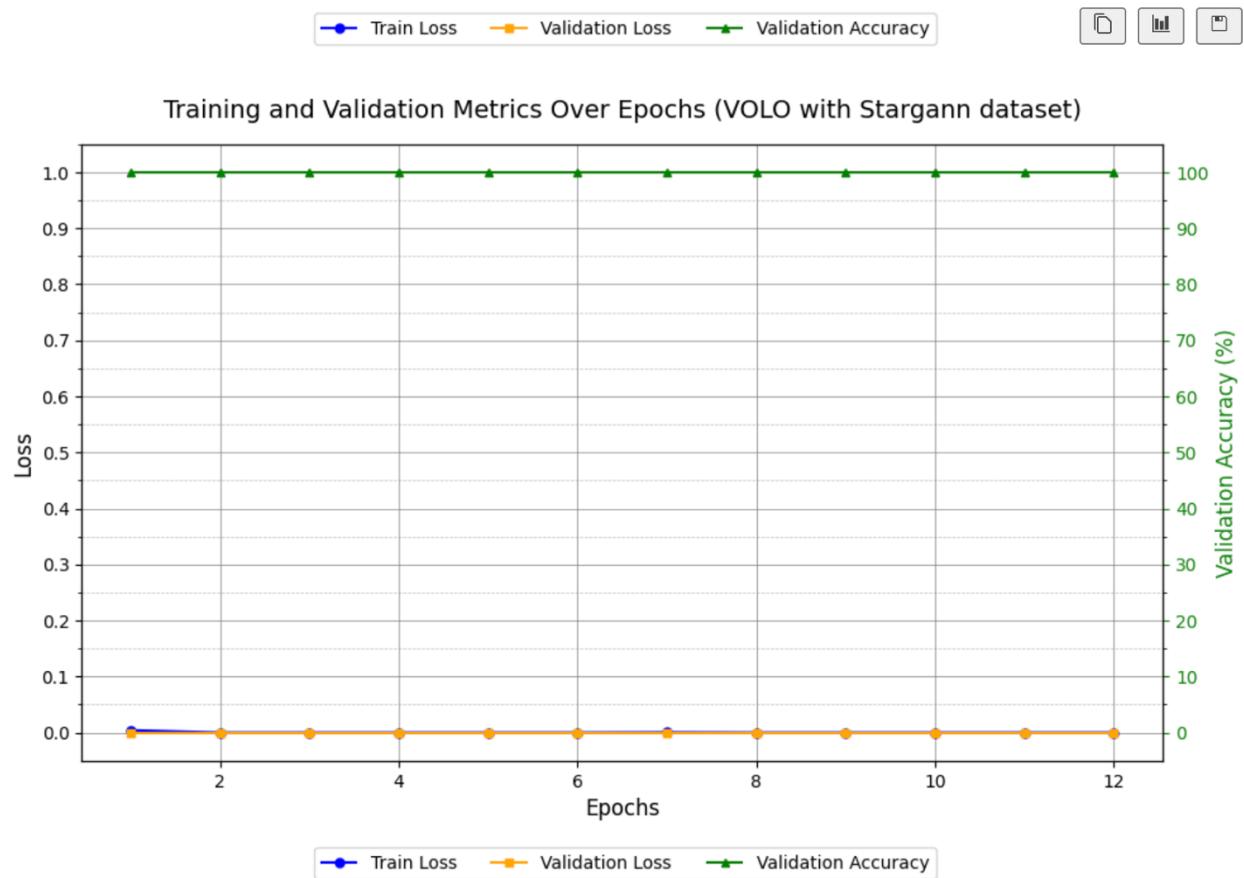
FTTQ DATASET



SELF CREATED DATASET



DFFD DATASET



STARGANN DATASET: Both models show bad results (even the fake and real images were very well created, both type of images were so different)

