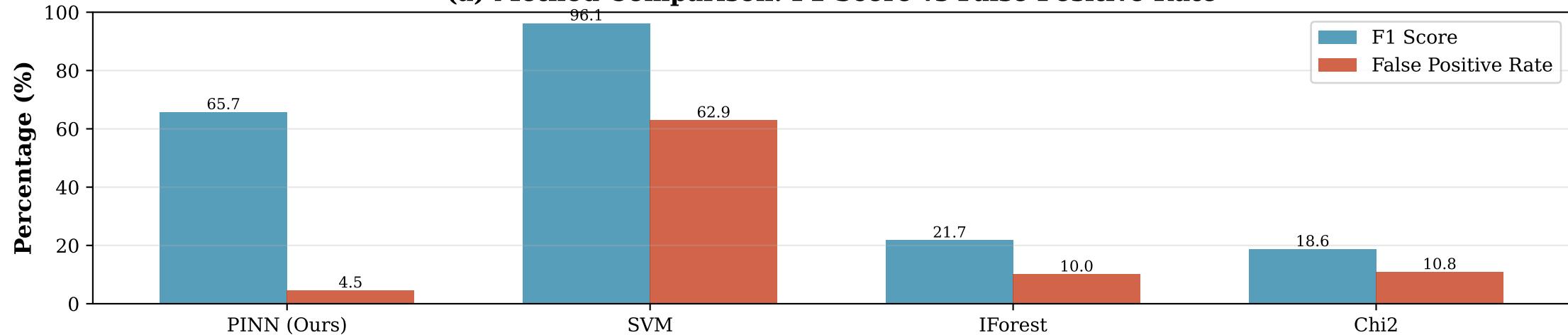
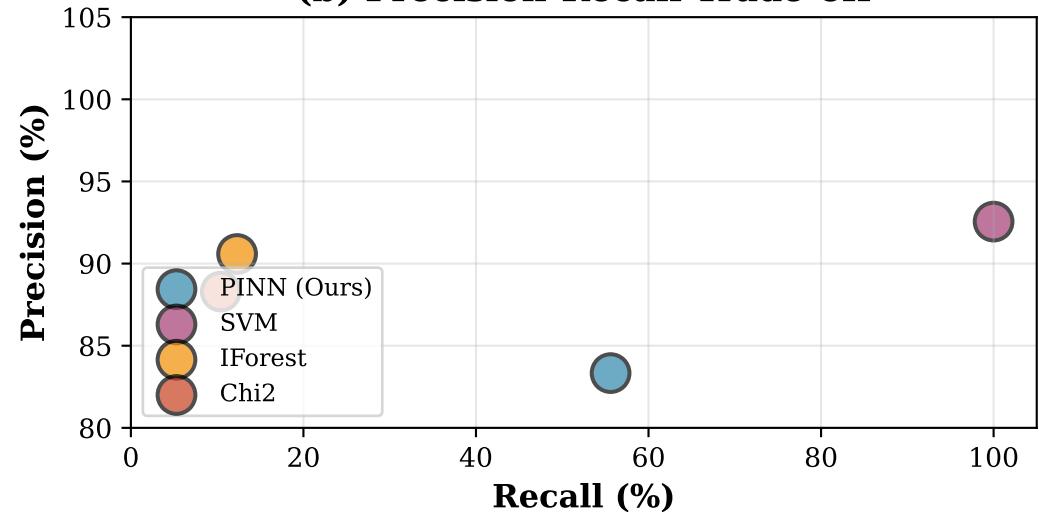


PINN-Based Fault Detection: Comprehensive Evaluation

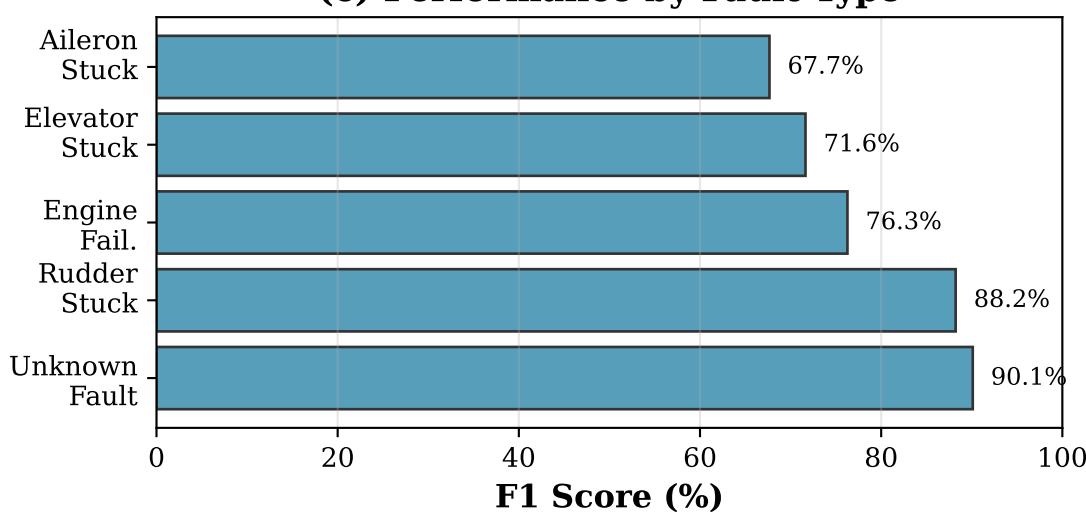
(a) Method Comparison: F1 Score vs False Positive Rate



(b) Precision-Recall Trade-off



(c) Performance by Fault Type



PINN Fault Detector - Key Results on ALFA Dataset (47 Flights, 5 Fault Types)

Overall Performance:

- F1 Score: 65.7%
- Precision: 83.3%
- Recall: 55.6%
- False Alarm Rate: 4.5% (BEST)

Comparison with Baselines:

- PINN achieves LOWEST false alarm rate (4.5% vs 62.9% for SVM)
- 2nd best F1 score (65.7% vs 96.1% for SVM)
- Balanced performance (not overfitting to high recall)
- Traditional methods (Chi2, IForest) perform poorly (F1 < 22%)

Conclusion: PINN detector provides the best balance between detection accuracy and false alarm rate, making it suitable for real-world deployment where false alarms are costly.