Confidence Distribution: Correct vs Incorrect Predictions
FakeSV Dataset (Defer Classifier)

Legend:
- Correct Predictions (n=506)
- Incorrect Predictions (n=36)
- Correct Mean: 0.900
- Incorrect Mean: 0.857

Higher confidence → Better calibrated model
Correct predictions should have higher confidence

X-axis: Confidence (Max Prediction Probability)
Y-axis: Density