Problems in Existing System and Potential Solutions

Limitation	Example System	Observed Shortcoming	Potential Solution(s)
1. No unified multi-format support	Deepware, Reality Defender	Only handle one media type	➢ Develop a **modular multimodal framework** that combines image, video, audio, and text analysis. ➢ Use **shared feature fusion layers** and **transformers** that handle multimodal inputs (e.g., CLIP, FLAVA, MMF). ➢ Incorporate **metadata and audio analysis** modules (e.g., speech forgery, text anomalies).
2. Real-time detection issues	DFDC top models	High GPU demand, not edge-optimized	➢ Implement **lightweight models** (e.g., MobileNetV3, TinyViT, DistilBERT for audio/text). ➢ Use **model compression** techniques: pruning, quantization. ➢ Deploy using **TensorRT**, **ONNX Runtime**, or **CoreML** for edge acceleration. ➢ Adopt **pipeline batching** and **frame skipping** for efficient video inference.
3. Limited generalization to new forgeries	FaceForensics++-based models	Fails on newer manipulations	➢ Apply **domain adaptation** and **meta-learning** to train models that generalize across datasets. ➢ Use **self-supervised learning** to extract features independent of specific forgery types. ➢ Train on **synthetic variations** using style-transfer GANs or diffusion models to simulate unseen manipulations.
4. Lack of explainability	Microsoft Video Authenticator	Black-box detection	➢ Integrate **Explainable AI tools**: Grad-CAM for CNNs,

			LIME/SHAP for tabular/audio models.
			➢ Provide **highlighted regions**, **confidence scores**, or **temporal tracebacks** in user interfaces.
			➢ Design **rule-based post-processors** to convert features into interpretable insights.
5. Adversarial vulnerability	CNN-based detectors	Easily fooled by noise	➢ Incorporate **adversarial training** (FGSM, PGD, DeepFool perturbations).
			➢ Use **robust loss functions** (TRADES, confidence-penalty, logit squeezing).
			➢ Add **input transformation defenses** (JPEG compression, bit-depth reduction).
			➢ Monitor **feature-space consistency** using ensemble voting or detection heuristics.
6. Dataset bias	Celeb-DF, FaceForensics++	Lack demographic diversity	➢ Curate **inclusive datasets** representing gender, ethnicity, lighting, compression, and language variance.
			➢ Use **federated learning** across global sources to build diverse, privacy-respecting models.
			➢ Monitor fairness with **bias evaluation metrics** (e.g., False Positive Rate per group).
7. No real-world deployment tools	Academic GitHub repos	No APIs or GUI	➢ Build **API-first systems** using Flask/FastAPI + Docker for modular services.
			➢ Develop **user-friendly dashboards** for journalists, forensic analysts, moderators.

			➤ Integrate with **cloud services** (GCP, AWS, Azure) for scalable usage. ➤ Offer **browser extensions or mobile apps** for public/media-facing workflows.