Drawbacks of Existing Systems:

1. **Lack of Unified Mutli-Format Support**
   - Most existing systems focus on detecting manipulation in a single media format—typically images or videos.
   - There is a notable absence of platforms capable of seamlessly analyzing images, videos, audio, and textual metadata together in one pipeline.
   - This limits the system's applicability in real-world environments where misinformation spans multiple content types.

   Example:

   I. Deepware Scanner (Mobile app) – Detects deepfake videos but does not handle images, audio, or text.
   II. Reality Defender – Offers browser-based image and video analysis, but not audio/text in one pipeline.

   Limitation - Each tool is domain-specific. There's no single system that can analyse all formats (image, video, audio, and text) in a unified, scalable platform for real-time media validation.

2. **Challenges in Real-Time Processing**
   - High-performing deep learning models often require substantial computational resources, making real-time deployment difficult.
   - This restricts use cases such as social media moderation, video conferencing, or live surveillance, where instant detection is critical.

   Example:

   Facebook's Deepfake Detection Challenge (DFDC) – Top models (like EfficientNet-based solutions) achieved high accuracy but required heavy computational resources (multi-GPU inference) and long processing times.

   Limitation – These models are impractical for deployment on mobile or edge devices or for detecting fake content during live streaming or video calls.

3. **Limited Generalization to Unseen Forgeries**
   - Detection models trained on specific datasets often fail to generalize well to newly generated or domain-specific fake content.

- As generative AI tools evolve, existing models become outdated unless continuously retrained.

Example:

FaceForensics++ Based Detectors – Most academic models trained on this dataset (e.g., XceptionNet) perform poorly on newer datasets like WildDeepfake or DeepfakeTIMIT.

Limitation – Models trained on older GANs or dataset-specific artifacts fail when confronted with new forgery styles, compression settings, or post-processing used in real-world apps like TikTok or Instagram Reels.

4. **Lack of Explainability**
   - Deep learning models tend to function as black boxes, offering little insight into their decision-making process.
   - In high-stakes fields such as journalism, forensics, and law, explainability is essential to ensure accountability and transparency.

Example:

Commercial Deepfake Detectors – While technically accurate, these systems don't provide **human-interpretable explanations** for detections. **(e.g., Microsoft Video Authenticator)**

Limitation – In sensitive applications (e.g., court evidence, media ethics reviews), explainability is essential for stakeholder trust, but these systems operate as black boxes.

5. **Susceptibility to Adversarial Attacks**
   - Subtle, imperceptible modifications to media inputs can mislead models, allowing fake content to bypass detection.
   - This poses a significant security risk, particularly in malicious or state-sponsored disinformation campaigns.

Example:

XceptionNet, VGGFace2, ResNet models – Researchers have shown that small perturbations, generated using methods like FGSM or PGD, can trick these detectors into classifying deepfakes as real.

Limitation - No robust defence against adversarial crafted media, even though attackers can automate such manipulations.

6. **Dataset Bias and Limited Diversity**
   - Public datasets are often biased towards specific demographics or conditions (e.g., celebrity faces, studio lighting), leading to inconsistent performance across diverse user groups and real-world scenarios.

   Example:

   Celeb-DF, FaceForensics++ - These datasets primarily feature Western celebrity faces under ideal lighting conditions.

   Limitation – Detectors trained on them perform worse on underrepresented ethnicities, varied facial structures, or cultural video content (e.g., regional YouTubers or Bollywood scenes).

7. **Lack of Real-World Deployment Interfaces**
   - Many systems remain in the prototype or academic research stage without user-friendly deployment interfaces such as APIs, web apps, or mobile tools, limiting their practical impact.

   Example:

   Many academic models published on **GitHub or PapersWithCode** (e.g., DeepFakeDetection, DeepFaceLab) are not available as APIs or packaged apps.

   Limitation - Lack of integration with content platforms, newsrooms, legal systems, or cloud service making these tools difficult to adopt outside research labs.