A Systematic Literature Review on Datasets for Deepfake Images in Smart Cities

ANAND M K
National Institute of Technology Karnataka, Surathkal
242CS008

Abstract

- Deepfake image datasets are critical for developing robust detection systems in smart cities.
- This review systematically examines:
  - Key datasets: DFDC, DeeperForensics, FaceForensics++, DeepfakeTIMIT, UADFV.
  - Deepfake generation techniques: GANs, Autoencoders, and hybrid models.
  - Preprocessing methods: Data augmentation, feature extraction, artifact analysis.
  - Detection methods: CNNs, RNNs, and hybrid models.
- Challenges: Dataset biases, generalization issues, real-time processing requirements.

Introduction

- **Deepfake Technology:**
  - Rapid advancements in AI and deep learning enable the creation of highly realistic synthetic images and videos.
  - Applications: Entertainment, education, accessibility.
  - Risks: Identity theft, misinformation, unauthorized access.
- **Smart Cities:**
  - Rely on IoT, AI, and big data analytics for urban infrastructure and services.
  - Vulnerabilities: Surveillance systems, authentication protocols, public information dissemination.
- **Need for Detection:**
  - High-quality datasets are essential for training and evaluating detection models.
  - Current datasets: DFDC, FaceForensics++, DeeperForensics-1.0, DF-TIMIT, UADFV.

Method

- **Literature Search Procedure:**
  - Databases: ScienceDirect, IEEE Xplore, ACM Digital Library.
  - Keywords: "deepfake detection," "image dataset," "smart city surveillance."
- **Research Problems:**
  - Focus on smart city security, deepfake detection, and dataset creation methodologies.
  - Inclusion criteria: Experimental studies, surveys, observational studies.
- **Search Strategy:**
  - Boolean operators: AND, OR.
  - Search string: "deepfake detection" AND "image dataset" AND "smart city surveillance."
- **Selection of Studies:**
  - Data extraction: Authors, publication year, dataset details, detection methods.
  - Analysis: Comparative evaluation of datasets and detection techniques.

DeepFake Detection Challenge (DFDC) Dataset

- **Overview:**
  - Largest and most diverse deepfake video dataset.
  - Contains over 100,000 video clips from 3,426 actors.
- **Generation Methods:**
  - Deepfake Autoencoder (DFAE): Shared encoder with two decoders.
  - MM/NN Face Swap: Custom morphable-mask model.
  - Neural Talking Heads (NTH): Meta-learning approach.
  - FSGAN: GAN-based face swapping and reenactment.
- **Augmentations:**
  - Distractors: Overlay objects like images, shapes, and text.
  - Augmenters: Geometric, color, and framerate transformations.
- **Detection Methods:**
  - Efficient ViT, Convolutional Cross ViT.
  - Pre-extraction of faces using MTCNN.

FaceForensics++ Dataset

- **Overview:**
  - Standardized benchmark for facial manipulation detection.
  - Contains 5,000 videos with real and manipulated content.
- **Manipulation Methods:**
  - FaceSwap: Graphics-based face region transfer.
  - DeepFakes: Neural network-based face replacement.
  - Face2Face: Facial reenactment.
  - NeuralTextures: Neural texture-based rendering.
- **Postprocessing:**
  - Simulates video quality degradation using H.264 compression.
  - Two levels: High-quality (HQ) and low-quality (LQ).
- **Detection Methods:**
  - Handcrafted features: Steganalysis with SVM.
  - Learned features: CNNs, XceptionNet, MesoInception-4.

DeeperForensics-1.0 Dataset

- **Overview:**
  - Largest face forgery detection dataset: 60,000 videos, 17.6 million frames.
  - Focus on real-world perturbations for robustness.
- **Generation Framework:**
  - DeepFake Variational Auto-Encoder (DF-VAE).
  - Disentangles structure (expression, pose) from appearance (texture, skin color).
  - Masked Adaptive Instance Normalization (MAdaIN) for style mismatch.
- **Diversity:**
  - 100 actors with varied genders, ages, skin tones, and nationalities.
  - Professional indoor setting with diverse lighting and camera perspectives.
- **Detection Methods:**
  - Image-level: Xception-Net.
  - Video-level: C3D, TSN.

DF-TIMIT Dataset

- **Overview:**
  - Created using GANs on VidTIMIT database.
  - Contains 640 videos (low and high quality).
- **Generation Process:**
  - Low-quality (LQ): 64x64 face regions, 200 frames at 4 fps.
  - High-quality (HQ): 128x128 face regions, 400 frames at 8 fps.
  - Blending techniques: CNN-based segmentation (LQ), landmark alignment (HQ).
- **Detection Methods:**
  - Lip-syncing detection: MFCCs, LSTM.
  - Image-based systems: PCA+LDA, IQM+SVM.

UADFV Dataset

- **Overview:**
  - Comprises 49 real and 49 Deepfake videos.
  - Average duration: 11.14 seconds, resolution: 294×500 pixels.
- **Detection Method:**
  - 3D head pose inconsistencies.
  - Features extracted using DLib and OpenFace2.
  - SVM classifiers for final classification.

Quantitative Comparison of Deepfake Datasets

Table: Quantitative comparison of various Deepfake datasets

Dataset	**Unique fake videos**	**Total videos**	**Total subject**
DF-TIMIT	640	960	43
UADFV	49	49	49
FF++ DF	4,000	5,000	?
DeeperForensics-1.0	1,000	60,000	100
DFDC	104,500	128,154	960

Conclusion

- **Dataset Evolution:**
  - Early datasets: UADFV, DF-TIMIT (small-scale, limited diversity).
  - Second-generation: FaceForensics++ (higher quality, ethical considerations).
  - Third-generation: DeeperForensics-1.0, DFDC (large-scale, diverse, robust).
- **Challenges:**
  - Dataset biases, generalization issues, real-time processing.
- **Future Directions:**
  - Improve dataset diversity and ethical compliance.
  - Develop scalable and real-time detection systems.