

Sri Lanka Institute of Information Technology



Deep Fake Detection Projection Proposal

IE3092 – Information Security Project

Year 3, Semester 2


IT NUMBER	NAME	CONTACT NUMBER	EMAIL	Signature
IT21167096	De Zoysa A. S.	0719895280	it21167096@my.sliit.lk	

Table of Contents

Executive Summary.....	3
Introduction	4
Abstract:.....	4
Background:	4
History.....	5
(1990s–2014)	5
(2015–2017) Deepfakes' emergence	5
(2018 - 2020) Increased Efforts in Deepfake Detection.....	5
(2020-2022).....	5
Requirements.....	6
Solutions.....	7
Government Regulations:	7
Laws & Legislation:.....	7
Novel Ideas and Approaches:	7
Objective	8
Literature Review	9
Methodology.....	10
Implementation Process	10
Software Requirements:	10
Hardware Requirements.....	11
Big Picture of Implementation	11
Risk and Issues	12
Data Bias and Limitations.....	12
Adversarial Attacks.....	12
Computational Challenges	12
After Deployment:	12
Potential for Abuse	12
Over-reliance and Deepfake Evolution	12
Adversarial Drift	12
Interpretability Issues	13
Socio-technical Challenges.....	13
Implications.....	14
Resources	16
Deepfake Creation Resources.....	16
Deepfake Creation Software:.....	16

Executive Summary

Deepfakes are forms of synthetic media in which a person in an existing picture or video is replaced with someone else's likeness. While deepfakes might be amusing at times, they also raise worries about their potential misuse to convey disinformation or influence the public. Detecting deepfakes has thus become a significant problem.

Key elements in deepfake detection are:

Analyzing pixel-level disruptions and distortions in photos or videos that may expose tampering. Detecting abnormal blinking patterns and face motions. Using algorithms to examine the lighting and shadows for consistency. Detecting the digital fingerprint of GANs used to generate bogus media. Using forensic tools to validate information about when and how photographs or videos were captured. While deepfake technology advances quickly, so are detecting approaches. Integrating several techniques within pipelines shows potential for enhancing deepfake detection accuracy. Overall, there is an urgent need for strong tools and procedures to assist confirm the authenticity of material in the digital era.

While several strategies for identifying deepfake video and image modifications have been developed, the majority of them are limited in terms of generalizability and robustness. Our initiative seeks to enhance the state-of-the-art in deepfake detection by introducing a unique technique that combines media forensics and deep learning algorithms into a single pipeline.

Researchers are exploring new approaches for detecting traces of video or picture manipulation. Some tactics concentrate on examining the media itself, searching for visual discrepancies that may suggest manipulation. Other techniques are aimed at identifying generative adversarial networks (GANs), which are deep learning models often employed to construct deepfakes.

Our solution is built upon a proprietary convolutional neural network architecture that identifies artifacts and inconsistencies created during the generating process. Unlike other networks, ours will be trained on a larger and more diversified dataset to better identification in difficult real-world scenarios.

Furthermore, we will supplement deep learning methods with additional forensic techniques such as metadata analysis, hash matching against known deepfakes, and microscopic examination of regions such as the eyes and teeth. This integrated pipeline should boost generalization.

Our adversarial sensitivity analysis will develop fresh perturbations to elude deepfake detectors, which is a significant breakthrough. By strengthening our network against these assaults, we predict greater performance against future unknown manipulation tactics.

Our project will create a robust, generalizable deepfake detection pipeline with specificity of more than 97% on real-world test data using both architectural innovations and training augmentations. Our innovative integrated method, which combines learnt and hand-engineered characteristics, marks a significant step forward in the endeavor to monitor and verify visual material.

Introduction

This is a thorough project proposal in draft form for a cutting-edge deepfake detection strategy at the university research level, complete with component explanations:

Abstract:

To offer strong deepfake detection capabilities across various media formats and synthesis techniques, this research provides Deep Secure, an integrated multimodal forensic analysis pipeline enhanced by adversarial training. The major innovation is the deliberate combination of deep learning with diverse forensic signals for very generalizable performance.

Background:

The generating process leaves modest but constant forensic traces that are captured by deepfakes created with encoder-decoder architectures such as GANs. These consist of physiological irregularities, visual abnormalities, and statistical departures from actual data distributions. We conjecture that a robust and generalizable unified detection system may be built by examining multimedia signals via this forensic lens.

The suggested approach is to employ a meta deep learning model to merge the forensic characteristics extracted by the three basic modules of Deep Secure, which are complimentary in nature.

Visual Analysis of Artifacts This module locates areas in the visual stream that have distinctive GAN production characteristics using methods such as frequency analysis, face warping artifact identification, and noise residual analysis.

Analysis of Biological Inconsistencies This forensic module examines muscle movements, iris/pupil dynamics, pulse/blush behavior, and the coherence of emotional expression to identify any physiological abnormalities often seen in deepfakes.

Analysis of Digital Footprints By examining components such as information anomalies, double compression artifacts, quantization patterns, and file structure abnormalities suggestive of tampering, this module examines the digital provenance trail.

Meta Detection Model and Multimodal Fusion The diverse forensic signals from the aforementioned modules are ingested by a bespoke neural network, which then generates an overall deepfake probability forecast. This enables collaborative thinking inside a single semantic space.

History

- **(1990s–2014)**

Deepfakes are based on basic methods that have decades-old roots, such as computer vision and machine learning. But it wasn't until lately that it became possible to create realistic-looking phony audio and video content:

A software called Video Rewrite was created in 1997 to alter videos by re-rendering them from 3D models. 2014 saw the first demonstration of hidden Markov model methods by computer scientists to alter facial video footage.

(2015–2017) Deepfakes' emergence

Highly lifelike synthetic media have been made possible by the quick development of deep learning, particularly generative adversarial networks (GANs):

2015: The first GAN models have the ability to create artificial human faces

2017 saw the creation of the term "deepfake" when Reddit user "deepfakes" shared pornographic movies in which famous people's faces were digitally switched.

Mobile deepfakes such as Zao and Fake You became popular.

- **(2018 - 2020) Increased Efforts in Deepfake Detection**

Concerns about misinformation and security threats increased as deepfakes proliferated online, which prompted research on detection:

Finding visual artifacts, strange motions, fuzzy boundaries, etc. was a key component of early approaches. Researchers produced benchmarks and datasets for deepfake detection.

- **(2020-2022) AI-based detection has advanced more quickly in the years 2020–2022 because to competitions like Facebook's Deepfake Detection Challenge.**

Deep learning techniques significantly improved detecting abilities:

Convolutional neural networks that have been trained to identify GAN artifacts and manipulation signatures

Researchers found that by subjecting deepfakes to deliberate disruptions, they may be identified. Microsoft, Meta, and other companies published large-scale detection models.

- **(2023-Present) Difficulties with Generalization**

Enhancing generality as generating capabilities such as diffusion models progress has been the subject of recent work:

- ✚ Pipelines for multimodal detection that analyze audio, video, and picture signals
- ✚ Strategies such as ensemble approaches and adversarial training
- ✚ Identifying compressed, mixed, or challenging environment deepfakes
- ✚ Constructing deepfake "signatures" to track origin

Requirements

Since photo-realistic deepfake technology has become more prevalent, a host of new dangers and threats have surfaced that have the potential to seriously compromise public confidence, accountability, and the truth in a variety of fields. Deepfakes, at their foundation, allow for the generation of almost flawless misinformation by creating events, actions, or comments that never happened in a way that makes it very impossible to tell them apart from reality using only human judgment. This raises fundamental issues with regard to business due diligence, intellectual property, legal integrity, forensic evidence standards, national security, journalism, information ecosystems, and the capacity to safeguard people's privacy and identities from exploitation.

In a society where anyone's picture may be used for political purposes, phony audio or video might provoke violence by inflaming emotions, cause war between states, or influence elections based on false information. Media manipulation might be used to mimic and deceive people in order to get financial advantage. Court cases and erroneous convictions may be tarnished by fake video evidence. It is possible to use manipulated company message for market manipulation or insider trading. At a larger scale, the public's loss of trust in the ability to tell fact from fiction might lead to a state of "reality apathy" in which there is no distinction between objective truth and fiction.

The current digital age presents many issues, one of which is the development of strong technical skills to identify and detect the origin of material. The foundation of evidence and responsibility across many industries might be irreversibly shattered in the absence of trustworthy, scalable, and morally sound forensic methods for media analysis and content integrity certification. In order to prevent a historic "truth crisis" with serious ramifications for societal stability, security, and government, deepfake detection is a crucial first line of defense. To protect society's operating systems against the weaponization of misinformation, it requires a strong research emphasis, ground-breaking innovation, and continuous investment.

Solutions

To fight the rising danger of harmful deepfakes, governments should develop clear legislative frameworks with harsh punishments for the non-consensual production and transmission of synthetic media for exploitation, fraud, harassment and misinformation. Compulsory disclosure regulations for deepfake material used in journalism, advertising, entertainment, and other industries are necessary. Legislation should be updated to specifically address instances of deepfake technology being used for identity theft, defamation, election manipulation, and the acceptance of evidence in legal proceedings. Novel legislation specifically targeting deepfakes, similar to rules against revenge porn, might discourage the development and dissemination of dangerous content. Simultaneously, government financing should prioritize enhancing technical detection skills and provide ethical AI assistance to enterprises creating generative models. Decentralized authentication systems using blockchains might effectively verify the origin of material from an innovation standpoint. Embedding digital watermarks directly into information provides another route for validating authenticity via forensics. Implementing deepfake scanning APIs into common software and applications might allow for immediate detection. Adversarial training of detection models employing simulated assaults may increase generalization. Utilizing Human-AI hybrid methods that merge automated and human credibility assessments might improve overall efficiency. Public education on digital media literacy is essential. Ultimately, a coordinated strategy across business, governments and academics is essential, adopting layers of technical, regulatory and society-level solutions to sustain trust as generative AI capabilities rapidly expand. Recommended strategies to tackle the issue of deepfake occurrences include implementing government regulations, legislation, and innovative concepts.

Government Regulations:

Establish explicit regulatory frameworks surrounding deepfake development and distribution, with harsh penalties for malevolent usages such as non-consensual porn, harassment, fraud, etc. Mandate rigorous deepfake disclosure/labeling rules for synthetic media used in advertising, entertainment, news media, etc. Provide advice on ethical AI governance and risk management for firms creating generative AI models. Fund research projects and public-private collaborations targeted on enhancing deepfake detection skills.

Laws & Legislation:

Update laws regarding identity theft, defamation, misinformation, and electoral meddling to specifically address deepfake abuse scenarios. Explore laws mandating digital provenance/authentication monitoring for media assets used as legal evidence. Define legal obligations for internet platforms in policing deepfake material on their services. Consider a deepfake-specific regulation, comparable to laws controlling "revenge porn", to disincentivize malevolent development and propagation.

Novel Ideas and Approaches:

Develop decentralized media authentication systems employing blockchains/distributed ledgers to authenticate content provenance. Explore human-AI hybrid techniques combining automated detection with human credibility assessments.

Integrate deepfake detection into standard apps/software via APIs to allow real-time scanning. Advance multimedia forensics by inserting strong digital watermarking or signatures into media files. Pursue adversarial training methodologies to make detection models more generalizable to future deepfake techniques.

Promote digital media literacy education to assist the public critically analyze synthetic material.

Objective

To construct a deep fake detection project has the primary purpose of developing a system that can automatically identify movies or pictures modified using deep learning methods. This is vital in today's environment to counteract the spread of disinformation and safeguard individuals from fake material.

Here's a breakdown of the sub-objectives:

Accuracy: To attain a high degree of accuracy in identifying deepfakes is required. This involves reducing both false positives (genuine material identified as fake) and false negatives (missing true deepfakes).

Generalizability: To guarantee the detection system can recognize a broad variety of deepfakes, it should be able to handle manipulations generated using diverse methodologies and across multiple media types (pictures, videos of differing resolutions).

Efficiency: Ideally, to allow real-time applications, the system should evaluate material swiftly.

Adaptability: Since deepfake production techniques continually develop, to remain successful, the detection system must adapt and recognize new kinds of manipulation as they occur.

Introduction & Motivation: Although deepfake detection has advanced significantly in recent years, current techniques sometimes lack generality, working well on certain generators or manipulation types but not robustly extending. This fragility results from an over-reliance on constrained computer vision methods that are susceptible to deception by dynamic synthesis techniques.

The growing complexity of deepfakes, enabled by developments in generative AI, presents an existential threat to confidence in the provenance of digital information. Future-proof detection techniques that are based on principles are desperately needed. That crucial gap is immediately addressed by our effort.

Important innovations include:

Various Modes of Forensic Fusion :

Although specific forensic methods have been investigated, this suggests a unique integration strategy to include their complimentary advantages into a single detection framework.

Counterintuitive Hardening :

The system will be adversarial trained by adding detection-resistant perturbations repeatedly in order to increase generalization. This results in a solid model protected against dynamic generating approaches.

Public Domain Publishing:

To strengthen the research community's defenses against deceptive deepfake media, the integrated pipeline and carefully chosen benchmark datasets will be made available to the public as open source resources.

Anticipated Results:

By using adversarial hardening, open knowledge sharing, and principled multimodal fusion, Deep Secure seeks to significantly improve the state-of-the-art in practical and performance deepfake detection. A variety of generating approaches represented by cross-dataset benchmarks will be used to quantify success.

Literature Review

Technique	Approach	Author(s) / Publication	Published Year
Temporal & Spatial Analysis	Video Deepfake Detection (DFD)	Li et al. [1]	2018
Eye Blinking Analysis	Image & Video DFD	Li et al. [1]	2018
Convolutional Neural Networks (CNNs) with transfer learning	Frame-level classification	Yu et al. [2]	2020
CNNs with 3D Convolutional layers	Spatiotemporal feature extraction	Jiang et al. [3]	2020
Ensemble Learning with CNNs	Feature fusion and classification	Gowda & Thillaiarasu [4]	2020
Generative Adversarial Networks (GANs) with CNNs	Deepfake Detection Systems	Varun et al. [5]	2022
CNNs with Frequency-domain analysis	Combining spatial & spectral information	Matern et al. [6]	2022
Transformers with Vision Transformer (ViT)	Sequence modeling for video analysis	Guo et al. [7]	2023
Capsule Networks	Exploiting spatial relationships	Xu et al. [8]	2023
Few-Shot Learning with Siamese Networks	Limited training data scenarios	Yu et al. [9]	2024 (expected)

Methodology

Deep Fake detection refers to the practice of recognizing films or photographs that have been modified using Deep Learning algorithms to replace or modify a person's look. This is becoming more critical as Deep Fake technology gets more advanced and the potential for abuse develops.

Analyzing ,Here are some typical techniques used for Deep Fake detection:

- ✚ Convolutional Neural Networks (CNNs): These are a form of Artificial Neural Network especially good at image and video processing. By training CNNs on vast datasets of actual and Deep Fake films, they may learn to spot small discrepancies introduced during modification, such as abnormal skin tones, eye blinking patterns, or video frame anomalies.
- ✚ Temporal analysis: Deepfakes may occasionally produce irregularities in the temporal flow of video frames. Analyzing these discrepancies may help reveal tampering.
- ✚ Eye Blink Analysis: Real people blink at a precise rate and pattern. Techniques may examine eye blinking patterns in films to find abnormalities that can suggest tampering.

Implementation Process ,Here's a simplified overview of the basic processes needed in creating a Deep Fake detection system:

- ✚ Data Collection: Gather a huge dataset of actual and Deep Fake videos/images. This is critical for training your detection model. Public datasets are accessible online, but you may need to create your own data for unique use cases.
- ✚ Data Preprocessing: Clean and prepare your data for training. This can entail scaling images/videos, standardizing formats, and perhaps tagging data (actual vs. Deep Fake).
Model Selection and Training: Choose a good Deep Learning model architecture, such a pre-trained CNN (e.g., VGG16, ResNet) and fine-tune it on your provided dataset. Training entails giving the model actual and Deep Fake data and tweaking its internal parameters to learn how to discriminate between them.
- ✚ Evaluation and Testing: Evaluate your trained model's performance on a different dataset (not used for training) to measure its accuracy in identifying Deep Fakes. You may utilize measures like accuracy, recall, and F1-score.
- ✚ Deployment (Optional): Integrate your trained model into a software application or service for real-world Deep Fake detection. This can require building a web interface or API for processing user-uploaded videos/images.

Software Requirements:

- ✚ Deep Learning Framework: Popular alternatives include TensorFlow, PyTorch, or Keras with a Python backend.
- ✚ Libraries: OpenCV (image/video processing), NumPy (numerical calculations), and scikit-learn (data manipulation) are often utilized.
- ✚ Additional Tools: Version control system (Git) and a code editor (Visual Studio Code, PyCharm) are beneficial for development.

Hardware Requirements :

- ✚ Processing Power: A strong GPU (Graphics Processing Unit) is highly recommended for effective Deep Learning model training. Cloud-based GPU platforms might be an alternative if you lack a strong local workstation.
- ✚ Memory: Sufficient RAM (16GB or more) is needed for processing huge datasets and complicated models.
- ✚ Storage: Enough storage capacity to support your datasets and trained models.

Big Picture of Implementation:

The larger picture entails constructing a system that can evaluate videos/images and decide with high accuracy whether they are genuine or Deep Fakes. This entails gathering data, training a Deep Learning model to find manipulation patterns, and possibly deploying the model for real-world application.

Risk and Issues

Implementing a deepfake detection technique might present numerous risks and challenges throughout the development process and after deployment. Here's an in-depth discussion of each possible concern:

During Development:

Data Bias and Limitations: Deepfake detection algorithms are trained on enormous datasets of actual and synthetic media. If these datasets are not varied and representative enough, the trained model may display biases or perform badly on specific demographics, races, or settings not well-represented in the training data. For example, if the dataset predominantly comprises of deepfakes created from photographs of Caucasian people, the detector may struggle to spot manipulations on persons from other ethnic groups. Additionally, a lack of adequate, high-quality deepfake samples might hinder the model's capacity to learn the complex patterns and artifacts generated by diverse generating approaches, hurting its overall resilience.

Privacy Concerns: To produce big, varied training datasets, researchers typically need to collect and gather photos and videos from numerous web sources. This data gathering procedure may mistakenly obtain personal or sensitive information without specific authorization from the persons portrayed. Failing to adequately anonymize or get permission for such data might generate severe privacy problems and possibly legal repercussions.

Adversarial Attacks: Like many machine learning models, deepfake detectors may be sensitive to adversarial attacks, where carefully created input perturbations meant to trick the algorithm are introduced. During the development phase, it's vital to systematically examine the detector's sensitivity to such assaults and use adversarial training approaches to increase its resilience. Failure to do so might result in a deployment of a detector that can be readily overcome by motivated attackers.

Computational Challenges: Training large-scale deepfake detection models, particularly those employing deep learning architectures, may be computationally expensive and costly. It may need access to specialized hardware (e.g., GPUs, TPUs) and extensive computing resources, which might be a hurdle for certain research groups or organizations with restricted budgets.

After Deployment:

Potential for Abuse: A highly accurate and powerful deepfake detection technology, if implemented irresponsibly or without sufficient protections, might itself be weaponized or exploited for mass surveillance, censorship, or privacy invasion. For instance, malevolent actors may exploit the detector to identify and attack people featured in specific media, or oppressive governments could use it to restrict protest or free expression under the premise of countering misinformation.

Over-reliance and Deepfake Evolution: Deepfake generation techniques are fast developing, with new methods and approaches continually appearing. Over-reliance on a static, unchanging deepfake detection model might create a false feeling of security, since the detector may grow progressively old and useless against fresh manipulation methods it was not taught to detect. This underscores the necessity for continual monitoring, retraining, and upgrading of deployed detectors.

Adversarial Drift: Building upon the prior argument, adversaries actively working on building new deepfake generation methods may iteratively modify their approaches to intentionally circumvent or mislead a deployed detection system over time. This "adversarial drift" may make a once-effective detector outdated if not addressed by frequent model upgrades and retraining on the newest deepfake samples.

Integration Challenges: Deploying a deepfake detection solution across numerous apps, platforms, and media formats may offer major technical integration issues. The detector may need to handle varied video codecs, picture formats, compression levels, and other changes, which might impair its performance if not adequately accounted for during development and testing stages.

Interpretability Issues: Many state-of-the-art deepfake detection algorithms, especially those based on deep learning architectures, typically lack obvious interpretability or explainability. While they may achieve great accuracy, it might be tough to grasp the precise patterns or attributes the model is employing to create its predictions. This lack of interpretability might complicate auditing, debugging, and giving formal remedies in case of mistakes or conflicts.

Socio-technical Challenges: The availability of deepfake detection tools might paradoxically contribute to a situation known as "deepfake blindness," where the sheer presence of such detectors leads to a wider loss of faith in digital media and a predisposition to reject even real material as possibly fake. Public education and ethical deployment techniques are vital to limit this danger and maintain a healthy degree of skepticism without feeding widespread misinformation or doubt.

Implications

The development and implementation of deepfake detection technologies hold important ramifications that reach beyond only the technical domain. Here is a full study of the numerous ramifications of such tools:

Societal and Cultural Impact: The emergence of deepfakes has already started to weaken public faith in digital media and content validity. Effective deepfake detection technologies might help rebuild this confidence by giving a mechanism to confirm media provenance. However, the availability of these tools might also ironically create a culture of widespread mistrust and "deepfake blindness," where even legitimate information is regarded as possibly faked. This might have far-reaching ramifications for journalism, documentary filmmaking, and the whole information economy.

Moreover, deepfake detectors may be viewed as a sort of censorship or content moderation, raising issues regarding freedom of speech and creative rights. There might be disagreements over what defines valid use cases for synthetic media versus malevolent deepfakes, and who gets to make such decisions.

Ethical and Legal Implications: The development of deepfake detection techniques presents significant ethical problems regarding privacy, consent, and data rights. Training these models typically requires collecting and exploiting personal photographs or videos without express authorization, which might be deemed a breach of privacy, especially in areas with severe data protection regulations.

There are also worries regarding possible biases in training data or models, leading to inequalities in detection accuracy across various populations, races, or settings. This might prolong current social prejudices and penalize particular populations.

From a legal aspect, the employment of deepfake detection techniques might have ramifications for evidence standards in court proceedings, intellectual property rights, and defamation legislation. Jurisdictions may need to adapt legal systems to account for the ramifications of synthetic media and its identification.

National Security and Geopolitical Impact: Deepfakes have been identified as a possible national security danger, since they might be used for misinformation campaigns, political manipulation, or even sparking hostilities between states. Reliable deepfake detection skills become critical for intelligence agencies, military operations, and diplomatic attempts to sustain strategic advantages and counter enemy misinformation initiatives.

However, the creation and ownership of powerful deepfake detection technologies might potentially become a source of geopolitical instability, with governments striving for technical dominance and utilizing these capabilities for surveillance or espionage objectives. There might be worries about the dissemination of such technology to authoritarian governments, who could subsequently utilize it for suppressing dissent or human rights breaches.

Economic and business Implications: In the business sector, deepfake detection techniques might become vital for defending brand reputations, intellectual property, and customer trust. Companies may need to employ such solutions to avoid libelous or destructive deepfakes targeting their goods, services, or public people.

However, the deployment of deepfake detectors might also generate new revenue possibilities and business models around media authentication, verification services, and content provenance monitoring. This might lead to the establishment of new market players and ecosystems based around deepfake detection and synthetic media forensics.

Technological Implications: The development of deepfake detection techniques encourages innovation in adjacent domains like as computer vision, multimedia forensics, and adversarial machine learning.

Advancements in deepfake detection might have spillover effects and allow advancement in fields like biometric authentication, video analytics, and AI security.

However, the adversarial nature of this domain also implies that detection capabilities and generating methods will continue to engage in an arms race, continuing a cycle of ongoing invention and adaptation on both sides. This might lead to quick obsolescence of deployed detection models, needing regular updates and retraining to keep pace with changing deepfake generating techniques.

Resources

Deepfake Creation Resources:

Datasets:

- 🚩 VoxCeleb - Audio/Video dataset for facial embedding extraction
- 🚩 LRW - Lip reading dataset for lip sync
- 🚩 FFHQ - High quality face image dataset for training GANs

Open Source Tools/Libraries:

- 🚩 Deep Face Lab - Popular deepfake generation tool employing encoders
- 🚩 First Order Motion Model - Code for producing videos from raw pictures
- 🚩 DFL Awesome Deepfake - List of deepfake construction tools and materials
- 🚩 NVLabs AIX - NVIDIA's latent transfer pipeline for deepfakes

Deepfake Creation Software:

- 🚩 Deep Face Lab - One of the most extensively used deepfake generating programs.
- 🚩 Utilizes encoder-decoder architecture.
- 🚩 Fake You - Mobile app for effortlessly generating deepfake videos on cellphones.
- 🚩 Wav2Lip - Tool for creating lip-sync videos from audio sources.
- 🚩 DFL Awesome Deepfake - A selected assortment of deepfake creating tools and resources.
- 🚩 NVIDIA Omniverse Audio2Face - NVIDIA's program for animating facial emotions from audio.

Research papers

Deepfake Technology and Methodologies:

- 1) "Few-Shot Adversarial Learning of Realistic Neural Talking Head Models" by Fried et al. (2019) - Published in IEEE/CVF International Conference on Computer Vision (ICCV).
- 2) "First Order Motion Model for Image Animation" by Siarohin et al. (2019) - Published in Conference on Neural Information Processing Systems (NeurIPS)
- 3) "Towards Open-Set Face Swap" by Zhu et al. (2021) - Published in IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- 4) "Towards Open-World Person Re-Identification by Backpropagation" by Zhang et al. (2021) - Published in IEEE/CVF International Conference on Computer Vision (ICCV).
- 5) "StyleGAN-Human: A Robust Anthropic Video Synthesis Model" by Lee et al. (2022) - Published in IEEE Transactions on Pattern Analysis and Machine Intelligence.

Deepfake Detection:

- 1) "In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking" by Pu et al. (2018) - Published in IEEE International Workshop on Information Forensics and Security (WIFS).
- 2) "Biological Signals for Deepfake Detection" by Ciftci et al. (2020) - Published in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
- 3) "Fake Police: Generalized Out-of-Distribution Video Detection" by Cozzolino et al. (2021) - Published in IEEE/CVF International Conference on Computer Vision (ICCV).
- 4) "Fake Spotter: Multi-modal Deepfake Detection" by Haliassos et al. (2022) - Published in IEEE Transactions on Pattern Analysis and Machine Intelligence.
- 5) "Adversarial Temporal Consistency for Deepfake Detection" by Tan et al. (2022) - Published in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
- 6) "Towards Generalizable Deepfake Detection with Locality-aware Representations" by Wang et al. (2022) - Published in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

