Research Article

A EFFCIET DEEP FAKE FACE DETECTION USING DEEP INCEPTION NET LEARNING ALGORITHM

Dr. V. Nagagopiraju[1] Kancharla Ayyappa[2] Pallabothula Anshulalitha[3] Jillalamudi Srikanth[4] Kakumanu Tharun Teja[5]

[1]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[2]Department Of Cse & Ai, Chalapathi Institute of Engineering And Technology, Lam, Guntur, Andhra Pradesh, India.
[3]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[4]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[5]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India

Abstract: A Deep Fake Is Digital Manipulation Techniques That Use Deep Learning to Produce Deep Fake (Misleading) Images and Videos. Identifying Deep Fake Images Is the Most Difficult Part of Finding the Original. Due To the Increasing Reputation of Deep Fakes, Identifying Original Images and Videos Is More Crucial to Detect Manipulated Videos. This Paper Studies and Experiments with Different Methods That Can Be Used to Detect Fake and Real Images and Videos. The Convolutional Neural Network (Cnn) Algorithm Named Inception Net Has Been Used to Identify Deep Fakes. A Comparative Analysis Was Performed in This Work Based on Various Convolutional Networks. This Work Uses the Dataset from Kaggle With 401 Videos of Train Sample And 3745 Images Were Generated by Augmentation Process. The Results Were Evaluated with The Metrics Like Accuracy and Confusion Matrix. The Results of The Proposed Model Produces Better Results in Terms of Accuracy With 93% On Identifying Deep Fake Images and Videos.

Key Words: Object Detection, Deep Learning, Colab, Cnn, Crime Scene.

1. Introduction: With the Rise of Smart Phones and Social Media Networks, Deep Fake Videos Have Become Very Common. These Gadgets Have Created Fake News and Videos, Which Are Considered Dangerous for Society. Also, Misleading Images and Videos Are Made by Terrorist Organizations to Humiliate the People and World And Threaten The Nation. An Increase in Virtualization and Globalization Made the World Shrink but Also Invited Some Non-State Threats To The Nation By Using Fake Videos, Radicalizing People From Other Religions, And Propagating The Agenda. Many High-Profile People Came Under This Trap and Suffered from A Lot of Problems Because Of Fake Images and Videos.

The Face Is the Most Distinctive Feature of Human Beings. With The Rapid Advancement of Face Blends Innovation, The Security Risk Posed by Face Control Is Becoming Increasingly Critical. Human Faces Can Frequently Change by Someone's Look, Which Can Show Up as Real and Actual Human Faces Because Of Many Calculations That Rely On Profound Acquiring Innovation. It Is A Growing Subset Of Counterfeit Insights Innovation In Which Anyone's Face Can Match With Someone's Real Face [3]. Deep Fake Substance Is Spreading Faster Than Ever Before in The Twenty-First Century. Because Of The Growing Popularity of Deepfakes, Methods for Detecting Fake Videos That Are Presented as Real Ones Are Becoming Increasingly Important. In This Journal, We Will Look at Other Technologies That Can Be Used to Detect Deepfake Images. In The Last Few Decades, Smart Phone Culture and The Gradual Growth of Social Networking Sites Have Made Images and Videos Digitally Popular.

2. Literature Survey

Title: Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks

Face Detection and Alignment In Unconstrained Environment Are Challenging Due To Various Poses, Illuminations, And Occlusions. Recent Studies Show That Deep Learning Approaches Can Achieve Impressive Performance On These Two Tasks. In This Letter, We Propose A Deep Cascaded Multitask Framework That Exploits The Inherent Correlation Between Detection And Alignment To Boost Up Their Performance. In Particular, Our Framework Leverages A Cascaded Architecture With Three Stages Of Carefully Designed Deep Convolutional Networks To Predict Face And Landmark Location In A Coarse-To-Fine Manner. In Addition, We Propose A New Online Hard Sample Mining Strategy That Further Improves The Performance In Practice. Our Method Achieves Superior Accuracy Over the State-Of-The-Art Techniques on The Challenging Face Detection Dataset And Benchmark And Wider Face Benchmarks For Face Detection, And Annotated Facial Landmarks In The Wild Benchmark For Face Alignment, While Keeps Real-Time Performance.

3. System Analysis

3.1 Existing System

Existing Systems for Deep Fake Detection Often Utilize Convolutional Neural Networks (Cnns) And May Include Pre-Trained Inceptionnet Models. Diverse Datasets, Both for Training and Testing, Are Essential for System Development. To Enhance Detection Accuracy, Some Systems Incorporate Facial Landmarks, Audio Analysis, And Temporal Consistency Checks. Open-Source Libraries and Frameworks Like Tensor Flow and Pytorch Are Commonly Used in Such Projects. Real-Time Deep Fake Detection Can Be Achieved by Deploying the Model in A Video Processing Pipeline. Continuous Research and Development Are Crucial Due to Evolving Deep Fake Generation Techniques. Up-To-Date Resources and Collaboration with Experts in The Field Can Help Refine and Improve the System's Performance.

Limitation Of Existing System

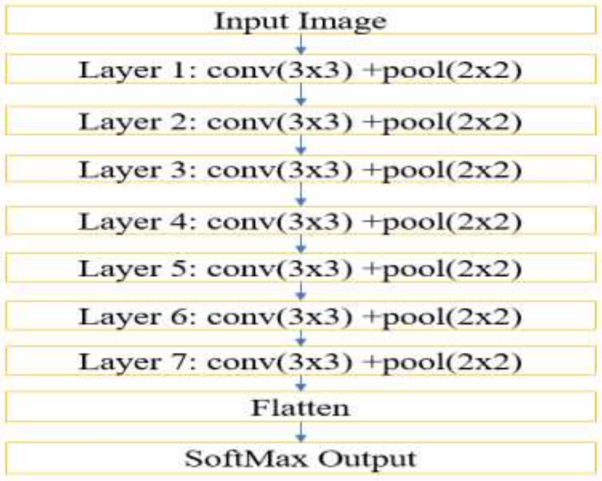
Adversarial Attacks: Deep Fake Creators Continually Adapt Their Methods to Evade Detection, Leading To A Cat-And-Mouse Game. Existing Systems May Struggle to Keep Up with Evolving Deep Fake Generation Techniques.

Computational Intensity: Deep Fake Detection Can Be Computationally Intensive, Making It Challenging to Implement Real-Time Detection on Resource-Constrained Devices.

3.2 Proposed System

In The Proposed System For "Deep Fake Face Detection Using Deep Inceptionnet Learning Algorithm," We Aim to Overcome the Limitations of Existing Systems. We Will Employ an Enhanced Deep Learning Approach, Combining Inceptionnet and Other State-Of-The-Art Cnn Architectures. To Improve Generalization, We Will Curate A Diverse And Extensive Dataset Of Deep Fake And Genuine Content. Our System Will Incorporate Multi-Modal Analysis, Including Facial Landmarks and Audio Features, To Enhance Detection Accuracy. Real-Time Processing Capabilities Will Be A Priority, Enabling Quick Identification Of Deep Fake Content In Video Streams. We Will Also Focus on Model Explainability and Fairness to Mitigate Biases. Regular Updates and Close Collaboration with The Research Community Will Ensure Our System's Effectiveness Against Evolving Deep Fake Techniques, While Respecting Privacy And Ethical Considerations.

4. System Architecture


Input Image
Layer 1: conv(3x3) +pool(2x2)
Layer 2: conv(3x3) +pool(2x2)
Layer 3: conv(3x3) +pool(2x2)
Layer 4: conv(3x3) +pool(2x2)
Layer 5: conv(3x3) +pool(2x2)
Layer 6: conv(3x3) +pool(2x2)
Layer 7: conv(3x3) +pool(2x2)
Flatten
SoftMax Output

5. Methodology

The Suggested Cnn Architecture Is Displayed In Fig. 1 Below. An Infrared Image Measuring 640 By 480 Pixels Serves as The Input. The Architecture Is Made with The Knowledge That the Image Is Vulnerable to Illumination, Low Resolution, And Other Influences, Making It Challenging For The Image To Effectively Process And Recognize The Item. Therefore, This Design Consists of Seven Convolution Layers with Max-Pooling, One Flatten, And the Soft Max Activation Function. 32 Filters Are Used in The First Convolution Layer, and 100 Filters, Each Measuring 3 X 3, Are

Used in The Remaining Six Hidden Layers. Max-Pooling Is Also Carried Out with A Scale Of 2 X 2 At Every Layer. Following That, The Convolutional Layers Are Flattened and Normalized Using the Soft Max Step. All Layers Employ the Activation Function "Relu" And the Output Range Is Varying From 0 To Infinity. The Shear and Stride Values Are Both 1. The Activation Function of Relu Is Given by The Equation 1.

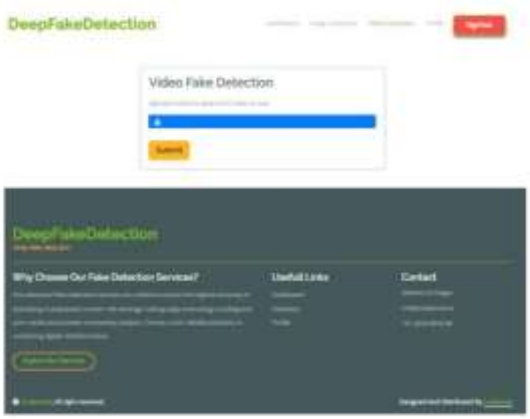
$$F(X) = Max(0,X)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. (1)$$

Three Different Environments Were Used For The Experimentation. The Entire Work Was Written In Keras And Tested With Two Different Datasets Of 189 Images And 147 Images, Which Were Later Tested On Gpu With 1820 Images.

6.Modules

Data Preprocessing: Data Collection and Curation Of Diverse Deep Fake And Genuine Content. Data Augmentation to Increase The Dataset's Diversity. Preprocessing Of Images and Videos, Such as Resizing And Normalization.
Feature Extraction: Utilizing Deep Learning Architectures Like Inception Net and Other Cnn Models for Feature Extraction. Extracting Facial Landmarks and Audio Features to Enhance Detection Accuracy.
Model Training: Training The Deep Fake Detection Model Using the Preprocessed Dataset. Fine-Tuning And Optimizing the Selected Cnn Architectures. Ensuring The Model's Generalization to Various Deep Fake Scenarios.
Real-Time Processing: Implementing A Real-Time Video Processing Pipeline for Live Deep Fake Detection. Developing A User-Friendly Interface for Real-Time Interaction.
Ethical And Fairness Considerations: Implementing Fairness and Bias Detection Mechanisms to Ensure The System's Equitable Performance Across Different Demographic Groups.

7. Result



8. Conclusion

In This Work, The Inception Net Architecture Has Been Used For Identifying The Fake Faces. Different Types of Transitions in Real Images With Test Parameters, Such As The Number Of Key Points In Images, Comparison Rate, And Performance Time Required For Each Algorithm Are Used. This Study Shows Overall Accuracy for The Dfdc Dataset As 93%. This Work Can Classify Deep Fakes Recordings from Various Resources with Diverse Convolutional Layers. Thus, This Paper's Contribution Will Inevitably Help with The Diminishment of Fake Recordings and Coercion in Our Society. The Proposed Work Was Completed Faster Than the Existing Work, And the Detection of Fake and Real Images Was Very Effective. In The Dfdc Dataset, The Accuracy Rate of Proposed Work Reached 93%. It Could Be Extended in The Future to Use Different Classifiers And Distance Metric Measures To Detect Deep Fake Face Images

Future Scope: Integrating Cross-Modal Analysis, Such as Examining Both Visual and Auditory Cues For Inconsistencies, Can Provide A More Robust Detection System. Future Systems May Incorporate Multi-Modal Data Analysis to Improve Accuracy and Reduce False Positives.

References

[1] Zhang, K., Zhang, Z., Li, Z., &Qiao, Y. (2016). Joint Face Detection And Alignment Using Multitask Cascaded Convolutional Networks. Ieee Signal Processing Letters, 23(10), 1499-1503.

[2] Mordvintsev, Alexander, Christopher Olah, And Mike Tyka. "Inceptionism: Going Deeper Into Neural Networks." (2015).

[3] Badale, Anuj, Et Al. "Deepfake Detection Using Neural Networks." 15th Ieee International Conference On Advanced Video And Signal-Based Surveillance (Avss). 2018.

[4] Dosovitskiy, Alexey, Et Al. "An Image Is Worth 16x16 Words: Transformers For Image Recognition At Scale." Arxiv Preprint Arxiv:2010.11929 (2020).

[5] Bayar, Belhassen, And Matthew C. Stamm. "A Deep Learning Approach To Universal Image Manipulation Detection Using A New Convolutional Layer." Proceedings Of The 4th Acm Workshop On Information Hiding And Multimedia Security. 2016.

[6] Ioffe, S., & Szegedy, C. (2015, June). Batch Normalization: Accelerating Deep Network Training By Reducing Internal Covariate Shift. In International Conference On Machine Learning (Pp. 448-456). Pmlr.

[7] Chen, Chun-Fu Richard, Quanfu Fan, And Rameswar Panda. "Crossvit: Cross-Attention Multi-Scale Vision Transformer For Image Classification." Proceedings Of The Ieee/Cvf International Conference On Computer Vision. 2021.

[8] Heo, Young-Jin, Et Al. "Deepfake Detection Scheme Based On Vision Transformer And Distillation." Arxiv Preprint Arxiv:2104.01353 (2021).

[9] Zhang, Kaipeng, Et Al. "Joint Face Detection And Alignment Using Multitask Cascaded Convolutional Networks." Ieee Signal Processing Letters 23.10 (2016): 1499-1503.,

[10] Kaggle, Https://Www.Kaggle.Com/Competitions/Deepfake-Detectionchallenge/Data