International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Application of Transfer Learning & Independent Estimation on Transformed Facial Recognition

Shivani Chaudhary[1*], Dr. Krishna Kumar Singh[2]

Submitted: 22/08/2023 **Revised**: 08/10/2023 **Accepted**: 22/10/2023

Abstract: Facial transformation is an interesting area that permits the recognition of a ideal facial to be moved to a marked facial by keeping the ideal facial's constraints. Consequently, the analysis utilizes deep learning and particular analysis for face exchange diagnosis. In the analysis, face exchange produces real images and time constraint orders. Humanoid detection could help in the preservation of face exchange occurrences. Humanoid helpers limit precision in integration faces. They may trust a picture that is a counterfeit. Ranks of feelings is a couple of Hamming-LUCBs which is deployed when we connect a picture. Choosing a substituted kind has a fifty percent fail rate. There are two types of deep learning approaches that is CNN-LSTM and RNN-LSTM that are compared here. Here in place of learning swapping faces we have executed the scheme standard parameters and errors. The trained and published data has different approaches. Consequently, they are being compared. The model's outcome has been estimated. The proportions of the facials are 0.2866 and 0.0416. Real/Fictitious equals 0.1107. Actual & false facials have parameters of 0.3702 and 0.4230. Amid 0.3742 and 0.1176. The study mentions differences between the true identities of Nirkin (unclear context) and AEGAN. Transfer learning is extensively utilized in this study. The "Best Face Swap Detection Photos in the World" are mentioned, involving around thousand real-life imageries for each model. The model's performance is evaluated by comparing two photos, where an imposter can be identified based on contextual cues. The mention of the model perceiving a human implies its ability to distinguish real faces from face-swapped ones.

Keywords: *Facial Transformation, Novel Technique, Transfer Learning, Independent Investigation, Procedure, Persons, Imageries*

1. Introduction

The facial swapping, which involves transferring a face from one image to another while maintaining the attributes of the target face. Face swapping has gained significant attention in computer vision and graphics research, and various techniques have been proposed to automate the process. There are two main approaches to face swapping: source-oriented and target-oriented. Origin basis methods work with the basis facial's imageries and blend it into the mark imageries through moving the mark facial's characteristics. These techniques rely over the posture & illumination of the basis imageries and may not accurately replicate the target's emotional state. On the other hand, expected-orientation methodologies straight modify the mark imageries' structures & are more flexible to variations in the basis facial.

While face swapping technology has advanced, there are concerns regarding its potential for misuse, particularly in the creation of "Deepfakes." Deepfakes are digitally edited images or videos created using deep learning systems to swap one person's face with another. This technology has raised concerns related to false information, such as fake news, hoaxes, and financial fraud. Researchers are working on developing detection methods to identify manipulated images and videos, drawing on techniques from mis spoof biometry and data analysed transfer learning. For swift countermeasures against deepfake technology. The detection of manipulated faces is an active area of research, and efforts are being made to develop techniques that can detect and identify fake images and videos. Subjective assessment techniques, involving human judgment, are being used in conjunction with deep learning methods to detect face swapping. Additionally, there are ongoing efforts to create community datasets & standards for evaluating the performance of diagnosis approaches. Recent research on face swapping, its applications, challenges, and concerns regarding its potential misuse. It emphasizes the need for advancements in detection methods to address the issue of manipulated media.

2. Proposed Methodology

A. Face swapping

In Muluye's approach (Muluye, 2020), a novel face "texture" is applied to a predictable 3Dimensional structure, which allows for approximations of section variable quantity viz. 3D angular representation & the resolution of the camera. This process is similar to Morphable Modeling, where all model parameters are optimized. By incorporating the new facial texture, the technique aims to enhance the

[1] *Research Scholar, Department of Computer Science & Engineering, Starex University, Binola, Gurgaon Haryana, India. Email: chshivani24@gmail.com*
[2] *Associate Professor, Symbiosis Centre for Information Technology, Pune, Maharashtra, India.*
Email: krishnakumar@scit.edu

realism of the generated images. On the other hand, Caruana and Vella (2020) propose a different approach that does not involve 3D reconstruction. Instead, their method focuses on finding a similar face in a large image library. The procedure includes placement a facial which resembles the involvement facial and determining which image from the library should be substituted. To enhance the realism of the swapped face, adjustments can be made to the lighting and color settings to better match the target image. However, this technique is limited to swapping one face for another and does not allow for more general face editing. Lutz & Bassett (2021): According to this paper, when working with referenced pictures that contain faces, the authors suggest replacing the faces with same methods & characters. They utilize an approach basis over triangular to regulate the orientation facial & contextual, making them appropriate the presence of the involvement facial and distorting the imagery. The facial regions of interest (ROIs) are identified utilizing a analysing process. Ashok et al. (2020): In this paper, Ashok et al. gives a brief representation of the basics of Poisson deletion. Poisson editing is a technique used for seamlessly blending or filling in regions in an image while preserving the overall structure and content. It involves generating a Laplacian for the image, and then solving the Poisson equation numerically to achieve smooth domain filling. This method can be applied to develop color imageries networks. The advancement of DL techniques, mostly in image processing, has led to the development of original facial-transformation approaches. These techniques involve the use of neural networks to convert and manipulate images, allowing for the swapping of faces between different individuals. Here is a summary of some of the techniques mentioned in your statement: Style Transfer: Ansari (2020) used style transfer techniques, considering face and posture as content and identity as style. By aligning faces and recording face landmarks from source and target photos, 3D face models were created. These models were then transformed and mixed using a fully convolutional neural network. Auto-encoders and GANs: Auto-encoders and reproductive argumentative systems (GANs) are working in Deep Fake, an increasingly automatic face-swapping technique. Ruelas et al. (2020) discovered a combined dormant interplanetary amid contribution and yield using automatic-encoders. By training two auto-encoders on the same encoder, common face features are learned. The unique decoders of each individual enable the generation of realistic representations from the latent space. Adversarial Loss: Sadu & Das (2020) use an auto-encoder to generate face images and utilize a convolutional neural network (CNN) as a discriminator. By applying an adversarial loss on switched faces, the quality of the generated images is improved. Hair and Face Swapping: Zhang & Doyle (2020) propose a method that swaps and replaces faces in the latent space, using hair and face regions. They employ a region-separative generative adversarial network (RS-GAN) to create a single face-swapped image. Video-basis Facial Spare: Wöhler et al. (2021) focus on moving picture-basis facial additional, where facials in cartridges are replaced with minimal human participation and low-cost equipment. These techniques demonstrate the progress in automated face-swapping methods, employing deep learning approaches such as style transfer, auto-encoders, GANs, adversarial training, and region-based manipulation. It's important to note that these techniques have ethical implications and can be misused for creating misleading or harmful content, such as deepfakes.

B. Detecting a forgery

Forgery detection refers to the process of identifying and determining whether a given object, document, or digital media has been altered or manipulated in an unauthorized or deceptive manner. Detecting forgeries is a critical task in various fields, including art authentication, document verification, forensic analysis, and digital media forensics. Several techniques and methodologies are employed to detect and expose different types of forgeries. Here are some common approaches. Visual inspection is often the first step in forgery detection. Experts with specialized knowledge in a particular domain carefully examine the object or document for any signs of tampering, such as inconsistencies in style, colour, texture, or handwriting. They compare the suspicious item with genuine examples or known standards to identify discrepancies. Forensic analysis involves using scientific techniques and specialized tools to examine physical or digital evidence. In art forgery detection, experts may employ various methods like infrared imaging, X-ray analysis, ultraviolet examination, or chemical analysis to study the materials, pigments, aging patterns, and underlying layers of a painting or artwork. Document examination focuses on the analysis of handwritten or printed documents, including signatures, contracts, or official papers. Forensic document examiners use techniques like handwriting analysis, ink and paper analysis, and magnification tools to identify irregularities, alterations, or discrepancies in the document that may indicate forgery. With the rise of digital media, detecting digital forgeries has become increasingly important. Digital forensics techniques involve analysing the metadata, file structures, compression artifacts, noise patterns, and other digital traces to identify manipulations or alterations in images, videos, or audio recordings. Advanced algorithms can detect inconsistencies in pixel-level details, compression artifacts, unnatural lighting, or perspective distortions. Statistical methods are often utilized to analyse large datasets or patterns to identify anomalies or suspicious patterns. This approach is commonly employed in financial fraud detection, where algorithms analyse transaction patterns, network connections, or behavioural data to detect fraudulent activities. Machine learning and artificial intelligence techniques have gained popularity in forgery detection. By training models on large datasets of genuine and forged samples, these algorithms can learn patterns and identify features that distinguish forgeries from authentic items. This approach is used in various domains, including image forensics, signature verification, and document authentication. It's important to note that forgery detection is an evolving field, and as forgers become more sophisticated, new techniques and countermeasures need to be developed. Collaboration between domain experts, forensic scientists, and technologists is crucial to stay ahead of forgers and ensure the accuracy and integrity of authentication processes.

3. Methods of Deep Learning

Deep learning is a subset of machine learning that focuses on training artificial neural networks with multiple layers to learn and extract meaningful representations from complex data. Here are some of the key methods used in deep learning:

Artificial Neural Networks (ANNs): ANNs are the foundation of deep learning. They consist of interconnected nodes (neurons) organized in layers. Each neuron applies a mathematical operation to its inputs and passes the result to the next layer. ANNs can have multiple hidden layers, enabling them to learn hierarchical representations.

Convolutional Neural Networks (CNNs): CNNs are primarily used for image and video processing tasks. They utilize specialized layers such as convolutional layers, pooling layers, and fully connected layers. Convolutional layers apply filters to input data, enabling the network to learn spatial hierarchies of features.

Recurrent Neural Networks (RNNs): RNNs are designed to process sequential data, such as text or time series data. They have a feedback mechanism that allows information to flow from previous steps to the current step, enabling them to capture temporal dependencies. Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs) are popular variants of RNNs.

Generative Adversarial Networks (GANs): GANs consist of two neural networks—a generator and a discriminator—competing against each other. The generator generates synthetic samples, while the discriminator tries to distinguish between real and fake samples. GANs have been successful in generating realistic images, videos, and other types of data.

Autoencoders: Autoencoders are unsupervised learning models that aim to learn compressed representations of data. They consist of an encoder that compresses the input data into a lower-dimensional representation and a decoder that reconstructs the original data from the compressed representation. Autoencoders can be used for tasks like data denoising, dimensionality reduction, and anomaly detection.

Reinforcement Learning: Although not exclusive to deep learning, reinforcement learning (RL) algorithms can be combined with deep neural networks to learn optimal actions in sequential decision-making problems. RL involves an agent interacting with an environment, receiving rewards or penalties based on its actions. Deep RL algorithms, such as Deep Q-Networks (DQNs) and Proximal Policy Optimization (PPO), leverage deep neural networks to approximate the value or policy functions. These are just a few of the prominent methods used in deep learning. The field is constantly evolving, and researchers are developing new architectures, optimization techniques, and regularization methods to improve performance and address various application domains.

4. Methods for Detecting Fake Faces

Detecting fake faces, often referred to as face spoofing or deepfake detection, is an important area of research in computer vision and artificial intelligence. Here are some methods commonly used for detecting fake faces:

1. Traditional Image Analysis: Traditional image analysis techniques can be used to identify inconsistencies or artifacts in images that may indicate tampering. These include checking for inconsistencies in lighting, reflections, and resolution, or detecting unnatural edges or boundaries in the image.
2. Micro-Expressions and Eye Blinking: One method to detect fake faces is by analyzing micro-expressions and eye blinking patterns. Fake faces generated through deepfake techniques may lack natural micro-expressions or exhibit abnormal blinking patterns. Analyzing these cues can help identify potential fakes.
3. Facial Movement Analysis: Facial movement analysis involves examining the facial landmarks and tracking the movement of various facial features, such as the mouth, eyes, and eyebrows. Fake faces generated using deepfake algorithms may have unnatural or inconsistent movements that differ from real faces.
4. Texture Analysis: Fake faces generated through deepfake methods often lack natural texture details or exhibit inconsistencies in texture. Analyzing the texture patterns and variations in different parts of the face can help identify potential fakes.
5. Generative Models Identification: Deepfake faces are often generated using generative models, such as Variational Autoencoders (VAEs) or Generative Adversarial Networks (GANs). Techniques that focus on identifying specific artifacts or signatures left by

these generative models can be employed to detect fake faces.

6. Deep Learning Approaches: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) can be trained on large datasets of real and fake face images to learn patterns and features that distinguish between them. These models can then be used to classify new faces as real or fake.

7. Temporal Consistency Analysis: Since deepfake videos are often created by stitching together different frames, analysing the temporal consistency of facial movements and expressions can help identify potential fake faces. Inconsistencies in the temporal patterns of facial movements can indicate manipulation.

8. Source Analysis: In some cases, it may be possible to analyse the source or origin of the image or video to determine its authenticity. This can involve examining metadata, reverse image searching, or investigating the credibility of the source.

5. Geometric Methods (Morphable Models)

Geometric methods, also known as morphable models, are mathematical and computational techniques used to represent and manipulate geometric shapes. These methods are widely employed in computer graphics, computer vision, and related fields to model and analyse 2D and 3D objects. Morphable models are typically built based on a set of example shapes or a training dataset. These shapes can represent various objects such as human faces, animals, buildings, or any other objects of interest. The goal is to create a model that can capture the variability and structure of the object class being modelled. One popular application of geometric methods is in the field of facial analysis and synthesis. Morphable face models (MFM) are constructed to represent the statistical variations in human faces. These models capture the shape and texture information of faces and can be used for tasks like face recognition, face tracking, facial expression analysis, and virtual face synthesis.

a. Reenactment and Puppeteering

Re-enactment and puppeteering are both forms of artistic expression and performance that involve bringing stories, events, or characters to life. While they share similarities, there are some key differences between the two. Re-enactment refers to the recreation of historical events or situations, often with a focus on accuracy and authenticity. It can be performed in various mediums, including theatre, film, or live events. Reenactors research and study the specific period, costumes, props, and actions associated with the event they are recreating. They aim to provide an immersive experience that educates and entertains the audience. Re-enactments are commonly used to depict historical battles, significant events, or daily life from a specific time period. Participants often wear period clothing, use authentic props, and follow specific scripts or improvisation guidelines. Re-enactments can range from small-scale local events to large-scale productions with hundreds or even thousands of participants. Puppeteering involves manipulating and animating puppets to tell stories or convey emotions. Puppets can be made of various materials, such as fabric, wood, foam, or even digital representations. Puppeteers control the movements, expressions, and voices of the puppets, giving them life and personality.

Puppeteering can be performed in different styles, including hand puppets, marionettes (string puppets), rod puppets, shadow puppets, or even animatronics. Puppeteers often work in teams, with some controlling the body movements, others handling facial expressions, and some providing the voice acting. Puppet shows can be performed on stage, in television or film productions, or as street performances.

b. Pipeline of Face-Swapping

Face-swapping, also known as face transfer or face replacement, is a technique used to replace or modify a person's face in an image or video with another person's face while preserving the original facial expressions and movements. The process typically involves several steps in a pipeline to achieve the desired result. Here is a general outline of the pipeline for face-swapping: The first step is to detect and locate the faces in the source image or video frame. Various face detection algorithms can be used for this purpose, such as Haar cascades, convolutional neural networks (CNN), or deep learning-based methods. Once the faces are detected, the next step is to align them properly. Face alignment techniques are employed to normalize the face regions by adjusting the position, scale, and orientation of the detected faces. This ensures that the subsequent steps can accurately manipulate the facial features. Facial Feature Extraction: After aligning the faces, the next step involves extracting the facial features from both the source and target images or videos. These features include landmarks, such as the positions of eyes, nose, mouth, and other characteristic points on the face. Various algorithms can be used for feature extraction, including shape models, active appearance models (AAM), or deep learning-based methods. Correspondence Estimation: In this step, the corresponding facial features between the source and target faces are determined. The goal is to establish point-to-point correspondences between the facial landmarks of both faces. This can be achieved using algorithms like facial landmark detection, optical flow, or shape matching techniques.

c. Network Architecture and Training

The network architecture refers to the structure and organization of the neural network model. There are various types of network architectures, each designed to solve specific types of problems. These are the simplest form of neural networks, where information flows in one direction, from input to output layers, without any loops or cycles. CNNs are widely used for image and video processing tasks. They have convolutional layers that can efficiently process grid-like input data, preserving spatial relationships. RNNs are designed to process sequential data, such as natural language or time series data. They have loops in their architecture that allow information to be persisted and shared across different time steps. LSTM is a type of RNN that addresses the vanishing gradient problem and can capture long-term dependencies in sequential data. It introduces memory cells and gates to control the flow of information. GANs consist of two components, a generator and a discriminator, which are trained together in a competitive manner. GANs are commonly used for generating synthetic data, such as images or text. The training process involves optimizing the parameters of a neural network model to minimize a specific objective function, typically a loss function.

d. Alignment and Stability of Landmarks

Alignment and stability of landmarks are important factors in various fields, including surveying, geodesy, mapping, and construction. Landmarks serve as reference points for measurements, positioning, and spatial analysis. Alignment refers to the spatial relationship between landmarks and their intended positions or coordinates. Proper alignment ensures accuracy and consistency in measurements and spatial referencing. Landmarks should be aligned according to a common geodetic reference system or coordinate system. This alignment is typically achieved using surveying techniques such as Global Navigation Satellite Systems (GNSS), total stations, or terrestrial laser scanners. Stability refers to the ability of landmarks to maintain their position over time. It is essential for long-term reliability and consistency of measurements and reference points. Factors that can affect the stability of landmarks include natural processes (e.g., tectonic movements, erosion), human activities (e.g., construction, excavation), and environmental factors (e.g., temperature, moisture). Unstable landmarks can introduce errors and inconsistencies in spatial data and analysis.

To ensure alignment and stability of landmarks, the following practices are typically employed:

- Geodetic Control Networks: Establishing geodetic control networks involves the precise measurement of reference points with known coordinates. These networks serve as a basis for aligning and referencing other landmarks. Control points are often established using high-precision surveying techniques and should be stable and well-distributed across the study area.
- Monitoring: Landmarks can be monitored periodically to detect any positional changes. Continuous GNSS monitoring or periodic re-surveys can be conducted to assess the stability of landmarks. Monitoring systems can alert users to any significant shifts in positions, allowing for necessary adjustments or investigations.
- Documentation: Comprehensive documentation of landmark positions, including their coordinates and associated metadata, is crucial. Proper documentation ensures traceability and allows for future reference and analysis. This documentation should be regularly updated as new measurements or surveys are conducted.
- Quality Control: Implementing quality control measures during surveying and mapping processes helps identify and mitigate errors. This includes ensuring proper instrument calibration, using redundant measurements, and conducting error analysis to assess the accuracy of measurements and alignments. By considering alignment and stability in the establishment and maintenance of landmarks, professionals can enhance the reliability and accuracy of spatial data, enabling more precise measurements, mapping, and analysis in various applications.

6. Comparing with Distinction in Manifold Groups

When the goal and basis imageries are scaled in location and appearance, photometric misalignment can cause visible seams, and differences in illumination can result in artifacts within the face. Edri et al. proposed a technique called multi-band blending that involves decomposing the source and target images into a Laplacian pyramid. Transitions between the images are smoothed close the cover's limits, but the illumination in the replicated area may still be incorrect. To address illumination issues, Ruelas et al. suggested recreating the low-frequency layers of the target's Laplacian pyramid and blending only the remaining detailed levels. This approach maintains the external facial contour. Kodali & Rekha introduced the use of GFC to match contrast between the source and target faces. The contrast coefficient is calculated based on GFC values and multiplied by each pixel, ensuring consistent contrast in the composited image. Cerda et al. conducted an ablation experiment and found that carefully selecting the blending mask can prevent the network's edges from introducing artifacts. They reduced the convex hull of the mask defined by outer facial landmarks.

7. Challenges

The issue of Deepfake forgery detection is indeed a challenging one, and the points you mentioned highlight some of the difficulties faced by researchers in this field. The proliferation of open-source face-swapping software and apps has made it easier for individuals to create Deepfake videos, leading to an increase in their presence on social media platforms. One of the main obstacles in developing effective detection techniques is the scarcity of good resolution transfer & real database that can be used for training and testing purposes. To build accurate detection algorithms, researchers require diverse and comprehensive datasets that encompass a wide range of Deepfake techniques and variations. Acquiring such datasets can be a significant hurdle due to the ethical and legal concerns associated with collecting and using real-world data. Moreover, the rapidly evolving nature of Deepfake technology makes it challenging to develop detection algorithms that can keep up with new forgery techniques. As detection techniques improve, creators of Deepfakes often adapt their methods to circumvent detection, leading to a constant cat-and-mouse game between detection and generation techniques.

Additionally, integrating detection algorithms with user interfaces that can effectively identify and filter Deepfake content is a complex task. The visual and audio manipulations involved in Deepfakes require sophisticated algorithms and processing power, making it difficult to implement real-time detection systems that can handle the scale and speed of social media platforms. Despite these challenges, researchers and organizations are actively working on developing and improving Deepfake detection techniques. They are exploring various approaches, including machine learning, computer vision, and forensic analysis, to detect inconsistencies and artifacts introduced by Deepfake manipulation. It is important to continue research efforts in this field to address the concerns posed by Deepfakes. Collaboration between academia, industry, and policymakers can help create a multi-faceted approach that combines technical solutions, legal frameworks, and media literacy programs to mitigate the potential harms associated with Deepfake technology.

Method

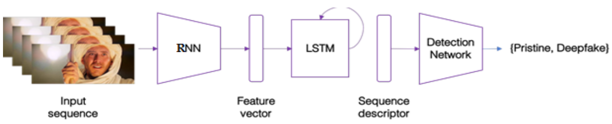

Fig. 1: Proposed flowchart

8. Dataset

Face swapping typically involves replacing one person's face in an image or video with another person's face while maintaining the facial expressions and movements. Auto-encoding is a technique used in machine learning for data compression and reconstruction. In the context of face swapping, an auto-encoder neural network can be trained to learn a compressed representation of a person's face, which can then be used to generate a new face with similar characteristics. To train an auto-encoder for face swapping, a large dataset of images of the same person is required. These images are often collected and prepared by academic researchers for their specific experiments. The dataset may go through processes such as removing blank images and duplicates, as well as cropping to eliminate unnecessary backdrops. The use of Google's picture API or other image sources can provide a wide range of celebrity images for training and testing the face swapping algorithm. Techniques like face detection and light detection can be employed to select appropriate images with diverse backdrops and lighting conditions. By using convolutional layers in neural networks, the algorithm can identify regions of interest, such as the face, before performing the face swapping. When selecting celebrities for the dataset, factors like gender and skin tone may be taken into consideration to ensure a diverse representation.

a. Sub-Subject Human

The study aims to rank photographs based on how accurately human volunteers can identify switched faces. The researchers argue that it's crucial to model uncertainty in the ratings, as the trust that an imagery is masked can be either unmistakable or uncertain. To gather valid ratings, multiple raters are involved, and each image pair is rated several times. Due to practical constraints, the study focuses on a subset of photos rather than the entire set. Specifically, the researchers manually selected 100 good quality transformed facial expressions & 200 original facial expressions by every method. This selection is made because randomly replaced faces are easily recognized, and an attacker would likely choose the best images for malicious use.

In the second technique employed, active learning is utilized to dynamically select the next image pair for approximate ranking.

b. Evaluations of websites

Website evaluations involve assessing and analyzing various aspects of a website to determine its effectiveness, user experience, and overall quality. These evaluations are typically conducted to identify strengths, weaknesses, and areas for improvement. Here are some common factors that are often considered during website evaluations: The visual design, layout, and aesthetics of a website play a significant role in creating a positive user experience. Evaluations may focus on factors such as overall visual appeal, consistency, use of colors and typography, and appropriate use of images. Website evaluations assess how easily users can navigate through the website and accomplish their intended tasks. Factors like clear and intuitive menus, logical page hierarchy, effective search functionality, and easy-to-understand links are evaluated to determine the website's usability. Evaluators assess the quality, relevance, and accuracy of the content on a website. They consider factors such as readability, grammar and spelling, comprehensiveness, and whether the content meets the needs and expectations of the target audience. With the increasing use of mobile devices, website evaluations often include an assessment of how well the website adapts to different screen sizes and devices. Compatibility with various browsers and operating systems is also considered. Evaluations may involve analyzing the website's performance, including page loading speed and overall responsiveness. A slow-loading website can negatively impact user experience and lead to higher bounce rates. Evaluations may consider how well the website adheres to accessibility guidelines, making sure it is usable for people with disabilities. Factors such as alternative text for images, proper heading structure, keyboard accessibility, and compatibility with screen readers are evaluated. For websites focused on conversions or specific actions, evaluations assess the effectiveness of CTAs, forms, and other elements that encourage users to take desired actions. Evaluations may also consider security measures, such as SSL encryption for data transmission, secure payment gateways, and proper handling of user data to ensure privacy.

9. Results and Discussion

Convolutional Neural Network -LSTM & Recurrent Neural Network-LSTM. In Fig. 2, you observed that the RNN-LSTM model illustrate a good percent contrast to the Convolutional Neural Network-LSTM model, indicating better performance.

Furthermore, you mentioned that the system recognized generalizing artifacts in the transformation as compared to the images within the training face expressions. This suggests that the models were able to identify common features or patterns associated with swapped faces rather than focusing on individual facial characteristics. Using default hyperparameters and a fixed number of epochs for the models. Additionally, you mentioned that you did not utilize an isolated validated database but instead integrated the different facial expressions by the approaches at the time of learning.

Table I. Comparison between Convolutional Neural Network-LSTM and Recurrent Neural Network-LSTM

Prototypical	**Training Precision %**	**Validation Precision %**	**Testing Precision %**
Convolutional Neural Network - LSTM	99.6	96.0	96
Recurrent Neural Network - LSTM	99.66	97.9	98.0

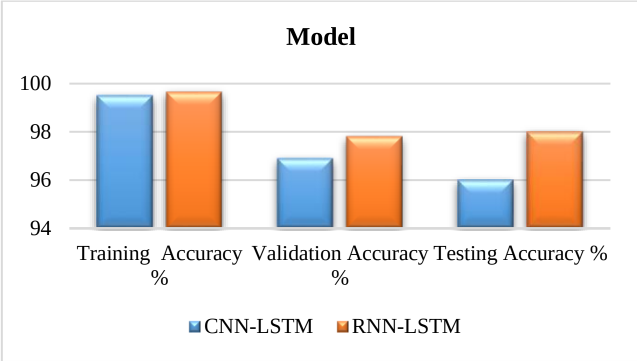

Fig.2. Comparison between CNN-LSTM and RNN-LSTM

The evaluation involves quantitative measurements, such as structural similarity (SSIM), as well as subjective user studies. The user studies aim to assess various aspects of the face-swapping results, including identity resemblance, head position, facial expression, and realism. The data used for these studies include samples from FaceSwap, Zhang et al. (2021), DeepFakes, and the researchers' own reshuffled face-swapping technique. To select a face that closely matches the description, it would be helpful to have access to the specific figures (Figure 3 to 5) and Table 2 that provide the frame comparisons and quantitative results. With that information, it would be easier to identify the most relevant face for further discussion. Additionally, the mention of Training Accuracy, Validation Accuracy, and Testing Accuracy suggests that the performance of the models (CNN-LSTM and RNN-LSTM) is being evaluated using these metrics for each frame. The techniques are also being compared to each other based on the approach described in Patidar & Bains (2021). This indicates that there were noticeable indicators of face swaps, but the quality of alterations varied. Issues such as contour mismatches, blurring, and unusual eye and mouth movements were reported. Viewers were able to detect high-grade modifications by examining behavioural cues like artificial expressions.

Table II. Comparison Different Frame

Model	**Training Accuracy %**	**Validation Accuracy %**	**Testing Accuracy %**
CNN-LSTM Frame-25	99.4	96.95	96.78
RNN-LSTM Frame-25	99.7	97.7	97.8

CNN-LSTM Frame-50	99.5	97	97.1
RNN-LSTM Frame-50	99.65	97.65	97.85
CNN-LSTM Frame-85	99.6	97.5	97.6
RNN-LSTM Frame-85	99.75	98.5	98.65

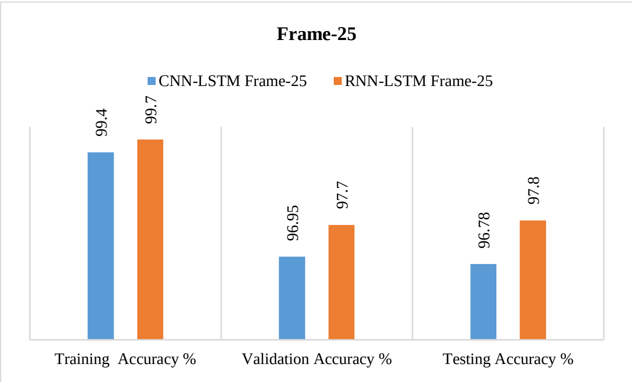

Fig 3: Comparison Frame-25 for between CNN-LSTM and RNN-LSTM

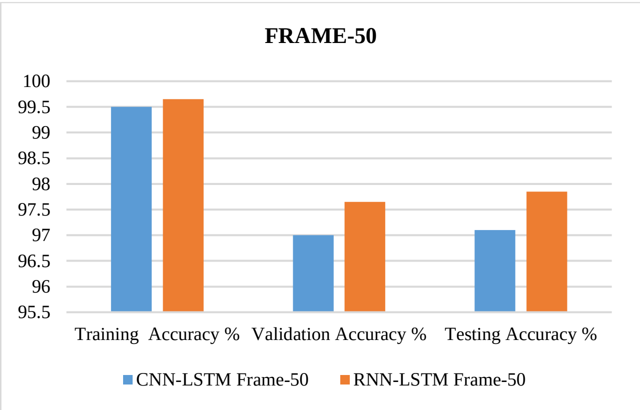

Fig.4. Comparison Frame-50 for between CNN-LSTM and RNN-LSTM

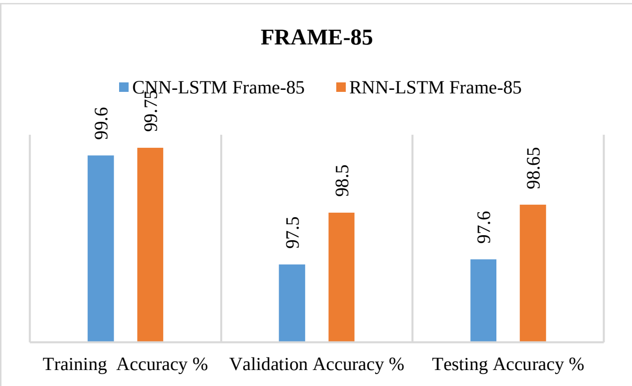

Fig. 5. Comparison Frame-85 for between CNN-LSTM and RNN-LSTM

The analysis of the neural network's outputs compared to human rankings suggests that there is a relationship between the model's ratings and the perceptions of humans, but it is not a perfect correlation. The correlation coefficients provide some insights into the similarity between the rankings, with Pearson correlation coefficients ranging from 0.2865 to 0.7896 and Spearman rank-order correlation coefficients ranging from 0.0415 to 0.7579. When comparing the ratings for natural and artificial faces, the correlation coefficients indicate a relatively weak relationship. For example, the Pearson correlation coefficient for natural faces is 0.2865, suggesting a positive but moderate correlation. On the other hand, the Pearson correlation coefficient for artificial faces is 0.0415, indicating a very weak positive correlation. Similarly, the Spearman rank-order correlation between real and fake faces is 0.1106, which also suggests a weak relationship.

The certainty level of the model, particularly in distinguishing phoney photographs created using the AE-GAN approach, shows less correlation with human rankings. The correlations between the model's certainty level and human ratings for AE-GAN generated images are lower compared to other techniques. This suggests that the model's rankings may diverge more from human perception when it comes to classifying these specific types of fake images.

It is important to note that the analysis mentions the separation between humans and the classifier in a high-dimensional space, implying that the model's representation of the data differs from human perception. However, the activation function's output can still be interpreted as a relative probability of class membership, indicating some similarity between the model's rankings and human rankings. Overall, while there is a relationship between the model's ratings and human rankings, it is not a perfect match. The model's rankings and human perception may diverge, especially when dealing with specific types of fake images. Further research and efforts are expected to improve the closeness in ranking between humans and the model.

10. Conclusion

In this manuscript a neural network basis solution to diagnose deep false videos. The model used here with huge amount of data illustrate very efficient outcomes. The new approach of associating feature extraction ability of CNN and classifying the ability of LSTM give improve outcomes as compared to existing models. This illustrates that the normal model that uses CNN with LSTM can be applicable to check if a video has been illustrated to any kind of manipulation. We understand that the work performed here gives effective results. The diagnosis model does not work for audio fake visuals. This problem will be resolved in our future work by exploring GAN for diagnosing the deep audio fake visuals.

References

[1] Ansari, N. (2020). *Dual Shot Face Detecting using Deep Learning*.

[2] Anwer, M., Ahmed, G., Akhunzada, A., & Siddiqui, S. (2021). *Intrusion Detection Using Deep Learning*. 1–6. https://doi.org/10.1109/ICECCME52200.2021.9590852

[3] Ashok, P., Vashisht, P., Kong, H.-O., Witt-Doerring, Y., Chu, J., Yan, Z., Oort, E. van, & Behounek, M. (2020). *Drill Bit Damage Assessment Using Image Analysis and Deep Learning as an Alternative to Traditional IADC Dull Grading*. https://doi.org/10.2118/201664-ms

[4] Bhavya, P., Sharmila, C., Sadhvi, Y., Prasanna, C., & Ganesan, V. (2021). *Pothole Detection Using Deep Learning* (pp. 233–243). https://doi.org/10.1007/978-981-16-1773-7_19

[5] Cárdenas, R., Beltrán, C., & Gutiérrez, J. (2019). Small Face Detection Using Deep Learning on Surveillance Videos. *International Journal of Machine Learning and Computing*, *9*, 189–194. https://doi.org/10.18178/ijmlc.2019.9.2.785

[6] Caruana, M., & Vella, J. G. (2020). *3D Facial Reconstruction from 2D Portrait Imagery*. https://doi.org/10.11610/isij.4724

[7] Cerda, B., Yuan, S., & Chen, L. (2021). *Phishing Detection using Deep Learning* (pp. 117–128). https://doi.org/10.1007/978-3-030-71017-0_9

[8] Chandrasekaran, A. (2021). Drowsy Face Detection using Deep Learning Algorithms. *International Journal for Research in Applied Science and Engineering Technology*, *9*, 2174–2179. https://doi.org/10.22214/ijraset.2021.35526

[9] Chen, D., Chen, Q., Wu, J., Yu, X., & Jia, T. (2019). Face Swapping: Realistic Image Synthesis Based on Facial Landmarks Alignment. *Mathematical Problems in Engineering*. https://doi.org/10.1155/2019/8902701

[10] Cui, S., Zhou, Y., Wang, Y., & Zhai, L. (2020). Fish Detection Using Deep Learning. *Applied Computational Intelligence and Soft Computing*, *2020*, 1–13. https://doi.org/10.1155/2020/3738108

[11] Ding, X., Raziei, Z., Larson, E. C., Olinick, E. V., Krueger, P., & Hahsler, M. (2020). Swapped face detection using deep learning and subjective assessment. *EURASIP Journal on Information Security*, *2020*(1), 6. https://doi.org/10.1186/s13635-020-00109-8

[12] Edri, E., Armon, N., Greenberg, E., Moshe-Tsurel, S., Lubotzky, D., Salzillo, T., Perelshtein, I., Tkachev, M., Girshevitz, O., & Shpaisman, H. (2021). Laser Printing of Multilayered Alternately Conducting and Insulating Microstructures. *ACS Applied Materials & Interfaces*, *13*(30), 36416–36425. https://doi.org/10.1021/acsami.1c06204

[13] Favole, F., Trocan, M., & Yilmaz, E. (2020, November 30). *Melanoma Detection Using Deep Learning*. https://doi.org/10.1007/978-3-030-63007-2_64

[14] Hou, X. (2018). An investigation of deep learning for image processing applications. *Undefined*. https://www.semanticscholar.org/paper/An-investigation-of-deep-learning-for-image-Hou/ebc1219ead03fe5720b1d118a5733f6e750ae9fe

[15] Kedar, E. (2021). Sarcasm Detection using Deep Learning. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*, 57–64. https://doi.org/10.17762/turcomat.v12i1S.1558

[16] Kodali, R., & Rekha, D. (2021). *Face Mask Detection Using Deep Learning*. 1–5. https://doi.org/10.1109/ICCCI50826.2021.9402670

[17] Korshunova, I., Shi, W., Dambre, J., & Theis, L. (2017). *Fast Face-Swap Using Convolutional Neural Networks*. 3697–3705. https://doi.org/10.1109/ICCV.2017.397

[18] Kumar, S. S. (2020). *Leveraging Big Data and Deep Learning for Economical Condition Assessment of Wastewater Pipelines*. https://doi.org/10.25394/PGS.12218432.V1

[19] Li, X., Lang, Y., Chen, Y., Mao, X., He, Y., Wang, S., Xue, H., & Lu, Q. (2020). Sharp Multiple Instance Learning for DeepFake Video Detection. *Proceedings of the 28th ACM International Conference on Multimedia*, 1864–1872. https://doi.org/10.1145/3394171.3414034

[20] Lutz, K., & Bassett, R. (2021). DeepFake Detection with Inconsistent Head Poses: Reproducibility and Analysis. *ArXiv*.

[21] Ma, T., Li, D., Wang, W., & Dong, J. (2021). *CFA-Net: Controllable Face Anonymization Network with Identity Representation Manipulation*.

[22] Maity, S., Das, P., Jha, K. K., & Sekhar Dutta, H. (2021). *Face Mask Detection Using Deep Learning* (pp. 495–509). https://doi.org/10.1007/978-981-16-3067-5_37

[23] Muluye, W. (2020). Deep Learning Algorithms and Frameworks for Deepfake Image and Video Detection: A Review. *International Journal of Engineering and Computer Science*, *9*(`10), 25199–25207. https://doi.org/10.18535/ijecs/v9i`10.4533

[24] Oh, B.-W. (2020). Map Detection using Deep Learning. *JOURNAL OF ADVANCED INFORMATION TECHNOLOGY AND CONVERGENCE*, *10*, 61–72. https://doi.org/10.14801/JAITC.2020.10.2.61

[25] Patidar, S., & Bains, I. (2021). *Intrusion Detection Using Deep Learning* (pp. 113–125). https://doi.org/10.1007/978-981-33-4305-4_10

[26] Prashanth, P., Vivek, K., Reddy, D., & Aruna, K. (2019). *Book Detection Using Deep Learning*. 1167–1169. https://doi.org/10.1109/ICCMC.2019.8819725

[27] Radenkovic, D., Keogh, S. B., & Maruthappu, M. (2019). Data science in modern evidence-based medicine. *Journal of the Royal Society of Medicine*. https://doi.org/10.1177/0141076819871055

[28] Ruelas, B., Dang, H., Nguyen, M., Dao, T., & Dhamo, D. (2020). *Truck Detection Using Deep Learning*.

[29] Sadu, C., & Das, P. (2020). Swapping Face Images Based on Augmented Facial Landmarks and Its Detection. *2020 IEEE REGION 10 CONFERENCE (TENCON)*. https://doi.org/10.1109/TENCON50793.2020.9293884

[30] Shekar, G., Revathy, S., & Goud, E. (2020). *Malaria Detection using Deep Learning*. 746–750. https://doi.org/10.1109/ICOEI48184.2020.9143023

[31] Singh, S. (2021). *Pneumonia Detection using Deep Learning*. 1–6. https://doi.org/10.1109/ICNTE51185.2021.9487731

[32] Thenmalar, R. (2020). Intrusion Detection using Deep Learning. *International Journal for Research in Applied Science and Engineering Technology*, *8*, 628–631. https://doi.org/10.22214/ijraset.2020.31521

[33] Verdhan, V. (2021). *Object Detection Using Deep Learning* (pp. 141–185). https://doi.org/10.1007/978-1-4842-6616-8_5

[34] Wöhler, L., Castillo, S., Zembaty, M., & Magnor, M. (2021). Towards Understanding Perceptual Differences between Genuine and Face-Swapped Videos. *CHI*. https://doi.org/10.1145/3411764.3445627

[35] Yadav, D., Renu, Ankita, & Anjum, I. (2020). *Accident Detection Using Deep Learning*. 232–235. https://doi.org/10.1109/ICACCCN51052.2020.9362808.

[36] Zhang, D., & Doyle, D. (2020). *Gate Detection Using Deep Learning*. 1–11. https://doi.org/10.1109/AERO47225.2020.9172619s

[37] Zhang, H., Ben, Y., Zhang, W., Chen, T., Yu, G., & Fu, B. (2021). Fine-grained Identity Preserving Landmark Synthesis for Face Reenactment. *ArXiv*.

[38] Luca Ferrari, Deep Learning Techniques for Natural Language Translation , Machine Learning Applications Conference Proceedings, Vol 2 2022.

[39] Kulkarni, A. P. ., & T. N., M. . (2023). Hybrid Cloud-Based Privacy Preserving Clustering as Service for Enterprise Big Data. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 146–156. https://doi.org/10.17762/ijritcc.v11i2s.6037

[40] Agrawal, S.A., Umbarkar, A.M., Sherie, N.P., Dharme, A.M., Dhabliya, D. Statistical study of mechanical properties for corn fiber with reinforced of polypropylene fiber matrix composite (2021) Materials Today: Proceedings, .