Forensic Science International: Synergy 6 (2023) 100309

Contents lists available at ScienceDirect

Forensic Science International: Synergy

journal homepage: www.sciencedirect.com/journal/forensic-science-international-synergy

Interpol review of forensic video analysis, 2019–2022

Zeno Geradts [a,b,*], Quinten Riphagen [a,c]

[a] Netherlands Forensic Institute, Laan van Ypenburg 6, 2497 GB, Den Haag, the Netherlands
[b] University of Amsterdam, Institute for Informatics, the Netherlands
[c] University of Twente, the Netherlands

1. Introduction

In this review, the most important developments are presented for the following general fields of expertise.

(1) Manipulation detection, and now in depth Deepfakes,
(2) biometric comparison (gait)
(3) Electric Network Frequency
(4) PRNU in Video
(5) Other fields of interest

The review focuses mostly on English literature, and which was available from public data sources from 2018 to 2022. Also we had to make a selection, since searches on video forensics in google scholar results in over 20.000 hits. For this reason the review is of highlights in fields, and there will be certainly papers and working groups and organizations that are missed.

2. Working groups and organizations

The development of forensic video analysis has several international working groups.

- **OSAC Digital/Multimedia Scientific Area Committee United States**
- **SWGDE Scientific Working Group on Digital Evidence** https://www.swgde.org/
- **ENFSI DIWG**: The European ENFSI Digital Imaging Working Group that is focused on methods, techniques, education and training. http://www.enfsi.org

3. American Academy of Forensic Science

Within the American Academy of Forensic Science the Digital and Multimedia Sciences Section works in this field.

Since 2003 each year a workshop was organized on Forensic Image and Video processing with handouts on the methods for face comparison, video restoration, 3D reconstruction, length measurement, photogrammetry and image processing. Also each year a scientific session was organized on this field. More information is available on: http://www.aafs.org.

4. ENFSI Forensic IT Working Group

The forensic IT working group of ENFSI deals with digital evidence as such. There exist some overlap with the Digital Imaging working group, and for that reason joint events are organized.

Since nearly all CCTV-systems are digital nowadays, often the question of handling the CCTV system itself is a question of digital evidence. Mobile phones, hard drives and other digital media should be handled in a secure way with proper forensic imaging software. The working group organizes training conferences each year. More information is available from http://www.enfsi.eu/.

5. Outline of this work

Since deepfake videos receive much attention, also due to the detection in forensic science, we will start with an literature review by Quinten Riphagen of the University of Twente who worked on a project at the Netherlands Forensic Institute on deepfakes. The number of references in deepfake forensics is around 3000 since 2018, so also here a selection is made.

The next field is biometrics based on gait analysis, and then an upcoming field is ENF in Video by the rolling shutter effect, and we will finalize the review with PRNU in Video.

6. Deepfakes

Deepfake detection papers will be divided by category and shown in Table 1. Missing accuracy's in the last columns are due to incomplete data in the article or an insignificant accuracy. The related definitions and further explanations needed to understand the paper will be provided in the sections below. Firstly, deepfake generation will be addressed. After which we will take a look at the different detection

* Corresponding author. Netherlands Forensic Institute, Laan van Ypenburg 6, 2497 GB, Den Haag, the Netherlands.
E-mail address: z.j.m.h.geradts@uva.nl (Z. Geradts).

https://doi.org/10.1016/j.fsisyn.2022.100309

Table 1
Overview of papers with category, and datasets.

Category	Paper	Contribution	Datasets accuracy
22*Frame-based	[1]	MesoNet	F2F(0.953)
	[2]	CNN(Visual artifacts)	DFDC(0.85
	[4]	Ensemble of CNN's(EfficientNet)	(AUC); DFD
			FF++(0.94
	[11]	Residual Noise + Transfer Learning	FF++(0.86
			DFDC(0.93
	[14]	Haralicks texture properties, SVM	FF++(¡0.8
			CelebDF(¡0
	[17]	Fourier Transform(POC)	None
	[19]	Adaptive Manipulation Traces Extraction	CelebDF + (combinati
	[21]	CNN(InceptionResNetV2)	CelebDF
			FF++@C2
	[22]	Mouth Region Analysis	CelebDF(0
			DFTIMIT(
	[26]	Single Class VAE OC-FakeDect	F2F(0.712)
			FS(0.861)
	[27]	Content feature extraction, trace feature extraction	Custom@c4
			F2F(0.858)
	[33]	PRNU + Image Cropping	FF++
	[35]	Transfer learning, SVM classifier	DFDC
	[37]	CNN(Spectogram image)	DFTIMIT
			(FF++(0.9)
	[39]	GAN Fingerprint	Own(up to
	[40]	Stacked Ensemble of Mod-els (DeepfakeStack)	FF++(0.99
	[45]	6 different CNN models with transfer learning	DFD + DF
			(Avg of 6 m
	[49]	Facial comparison (SVM)	CelebDF(0
	[50]	Modified AlexNet	FF++(0.87
			CelebDF(0
	[52]	Image Saliency CNN(ResNet18)	F2F(0.975)
			FS(0.957)
	[53]	Manipulation Classification Network	Own
	[55]	Multi-Layer Fusion Network	FF++@40c
			FF++@23c
	[56]	Multi-task CNN, 3D-ResNext	DFDC(0.97
			FF++(0.92
7*Temporal	[3]	Optical Flow, CNN	FF++(0.81
	[6]	Optical Flow Fields	FF++@40c
	[12]	Discrete Fourrier Transform + SVM	CelebDF
	[23]	Eye-blinking GAN	Own datas
	3 [28]	Sharp-Multi Instance Learning	CelebDF (0
			FF++(0.97
			DFDC(0.85
	[42]	FaceNetLSTM	FF++
	[54]	Time Series(LSTM)	FF++@C4
7*Spatiotemporal	[10]	Convolution Latent Representations combined with Bidirectional Recurrent	FF++, CelebDF

Structures(LSTM + CNN + RNN).
Structures(LSTM + CNN + RNN).

methods available to forensic experts to detect deepfakes and how they work. After which, the datasets used to train the models for detec-tion and benchmark the detection methods. Following that, various methods other than detection will be discussed. In addition to the chapter on detection methods, an overview consisting of all papers covering detection methods and the method they cover is provided at Table 1.

6.1. Deepfake generation

Deepfakes are most commonly created using Generative Adversarial Networks.

[16] provides the first Mathematical model of the generative adversarial net-works (GAN) that most Deepfake technology is currently based on. These networks are capable of making nearly undetectable generated media such as images, video and audio [16]. explains the way these networks create these undetectable media is by pitting a generative model against an adversarial dis-criminatory model whose responsibility is to determine whether a given sample is from the training model or from the generative model. The generative model learns from the discriminatory model by which samples passed through detec-tion and which were detected. In the following iterations, the features from the undetected samples are kept and expanded upon. This process is repeated for many generations until the discriminatory model can no longer accurately predict whether a given sample is real or generated. An example of the network structure can be seen in Fig. 1.

[16] explains this with an analogy of counterfeiters vs the police. The coun-terfeiters are trying to create the most realistic counterfeit money while the police keep working on new measures to detect counterfeit money. Competition in this field leads to nearly undetectable counterfeit currency. The same process applies in the creation of GAN generated deepfakes. Which makes detection by software measures a difficult problem.

6.2. Deepfake detection

The problem of detecting deepfakes in the research community is seen as a bi-nary classification problem, a video can either be classified as real or fake. Since classifiers need features to discriminate input data as either real or fake, deep learning methods are often used for as they offer automated feature extraction which is unrivaled in speed and accuracy to other methods. In forensic science, the problem cannot be approached in such a binary way, since the classification of a certain video needs to be explained. The forensic expert does not pass a.

Verdict on the authenticity of evidence, but explains using likelihood ratio's and analysis from the models what the chances are of the video being authentic. The verdict of whether the evidence holds up is then left up to the judge [38]. The full detection pipeline often consists of facial extraction from video frames, running the faces through a deep neural network to extract the features and train the model, and then classifying

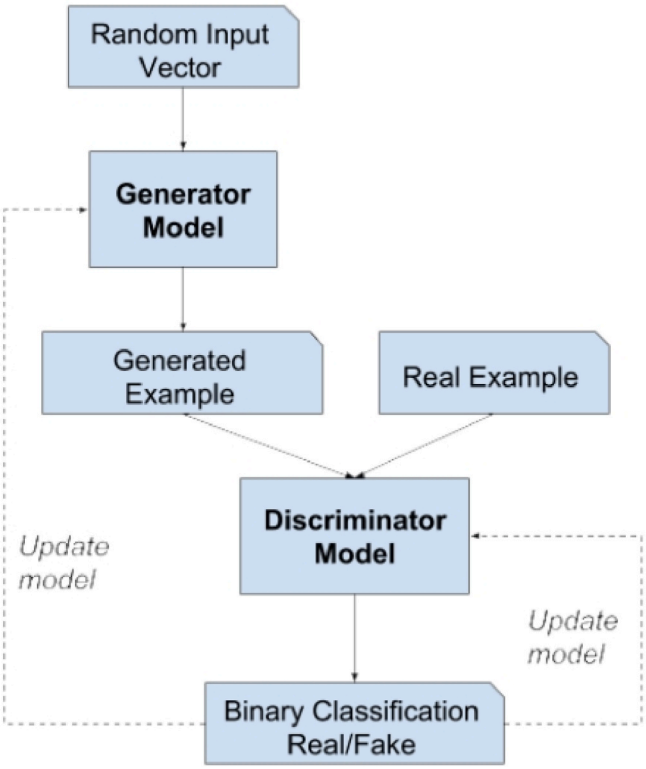

Fig. 1. Example of a GAN. Retrieved from [5].

the frames based on the extracted features using a binary classifier. An example of this structure can be seen in Fig. 2. In the subsections below we will discuss the different elements that most deep-fake detection models consist of.

6.2.1. Frame-based & temporal

Deepfake detection methods are divided into frame-based and temporal meth-ods. Frame-based detection methods detect deepfakes based on singular frames of videos and do not consider the temporal relationship between multiple consec-utive frames [36]. While temporal methods do consider the relationship between multiple consecutive frames of a single video. The frame-based are often more efficient as less frames are needed from a video to classify it as fake or real, but are often less accurate than temporal methods which consider more information when classifying videos. However, some frame-based methods outperform some temporal methods, in general though temporal methods will yield a higher clas-sification accuracy. Some models attempt to analyze both the temporal and the frame-based information hidden in the frames, which we will categorize as spatio-temporal methods [36].

When choosing between temporal methods and frame-based methods a con-sideration should be taken as to which goal the model needs to accomplish. The trade-off between efficiency and accuracy has to be taken into account. If the model needs the highest possible accuracy, then temporal methods have to be chosen. If however the model needs to run on lower quality hardware, such as mobile devices, and accuracy is less important, then an efficient frame-based method will suffice.

6.2.2. Loss functions

A loss function is essentially a measure of how well your algorithms models to the dataset. The output of the loss function when creating an algorithm can tell you whether the improvements you made to your model actually improve the model. A lower value representing an improvement of the algorithm on the same model.

[10] compare the use of two different loss functions and their effects on the detection accuracy. In particular, the authors compare the Cross-entropy (CE), the Kullback-Leibler (KL) divergence (relative entropy), and an ensemble (EN) of both functions. The authors report that the KL loss function outperforms the EN loss function on the face forensics dataset. For the CelebDF dataset, the EN loss function outperformed the KL loss function. Showing that different loss functions can be useful in different situations.

6.2.3. Artifacts

Most frame-based detection methods are based on so-called artifacts left by the GAN when generating the deepfake. These artifacts are basically traces of manipulation that can be extracted as features by the model. Some methods focus on specific artifacts while others take a more general approach.

[53] proposes a method to classify the various types of manipulations that have occurred on images. The manipulation classification network (MCNet) exploits multi-domain features to extract multiple features from the frequency, spatial, and compression domains [53]. This network is achieved through a multi-stream structure that and can detect image manipulation by analysing the fused features of multiple domains. The network can classify image manipu-lations and jointly considers compression and manipulations [53]. [53] describes a total of 6 types of commonly used image manipulations, which are image blurring, noise addition, contrast change, image morphing, image resampling, and JPEG compression. The network is composed of two different training networks; the Visual Artifact Network (VANet) and the compression artifact network (CANet). The features learned from these networks are then *trans*-ferred using transfer learning to the ensemble network which fuses the features into a single classifier [53].

[15] proposes a method based on the Ghost effect, which is a known effect in image forensics based on JPEG compression characteristics. The splicing of an image with different compression parameters when compressing the image again using the original images parameters, will result in a noticeable ghost effect in the spliced parts when image data is represented in terms of quantisized spectral DCT-coefficients [15]. The authors attempt to recreate this effect on video frames.

[27] aim to combine the fake face forensics with the more general fake image forensics, which exhibit different types of artifacts for the model to discriminate upon. Image forensics is mainly concerned with whether a certain image was manipulated or not, this includes resampling, resizing, compression, etc. [27]. While fake face forensics is mostly concerned with whether faces were manipu-lated [27]. propose a deepfake detection framework that combines these two types of image forensics. Combining a content feature extractor (CFE) and a trace feature extractor (TFE) in a single detection pipeline and aggregating the extracted features from both models. The CFE extracts facial content features such as facial tones, eyes, wrinkles etc. While the TFE receives input without the facial content of the image, which leaves it only trace information about the image such as minor texture differences and contour of the face [27]. The TFE then extracts features from an image with less information than the CFE and tries to find mostly general tampering evidence.

[52] tries to detect deepfake through image salience methods. Image saliency describes the things that are most noticeable in an image and can be measured through the difference in texture depth and pixels in an image. By enhancing the images through a guided filter, the authors enhance the texture details of fake images, making the difference between real images and fake one larger. Saliency detection is essentially a visual attention mechanism which recognizes the most noticeable parts of an image [52]. Manipulated images often have a noticeable different saliency in the manipulated regions of the image. The image saliency is represented in grayscale so we can visually see the difference between images.

[14] propose a novel method of detecting deepfakes based on Haralick's tex-ture properties. Specifically, the authors try to find anomalies in the properties of greyscale values in manipulated images. Additionally, the authors use focus measures which can be used to detect the "blurriness" of a certain image as an tool to increase the model's accuracy. Contrastingly to other deepfake de-tection methods, the authors decided to forgo using deep learning to extract features from the images, as the inner workings of neural networks can often be a black box to

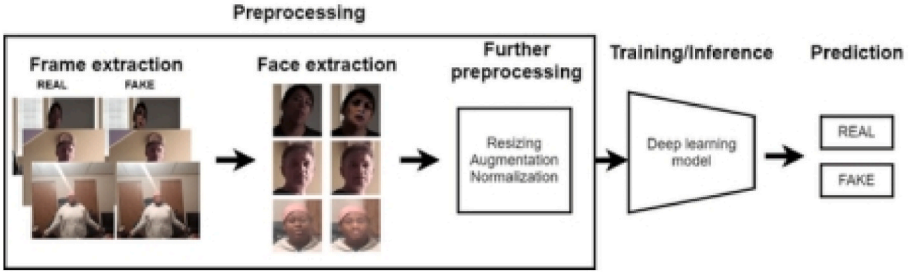

Fig. 2. Generic Deepfake detection pipeline. Retrieved from [9].

human observers and do not provide an explanation as to how the network decided on which features to extract. This method did not reach the desired results, after which [14] decided to use deep learning anyway.

[19] proposes an adaptive manipulation traces extraction network (AMTEN) to pre-process image content and filter out image content, highlighting only ma-nipulation traces. The authors say that this will improve the quality of classifi-cation of Convolutional neural networks since the image analysis will now only be focused on manipulation traces instead of all of the images content. The main reason they mention is that most detection methods do not take into ac-count the types of post processing such as compression, scaling, gaussian blur, etc. an image or video has received and will thus not take into consideration the difference between raw and processed images. Making it harder for networks to detect fake images if they are heavily processed. To combat this, the net-work utilizes adaptive convolutional layers to predict the manipulation traces in image content, which following layers can then use to maximize manipulation artifacts [19]. During back-propagation, the weights in neurons analyzing the manipulation traces are updated adaptively, which can then be used in subse-quent layers.

[11] propose a deepfake detection method based on residual noise, the dif-ference between the original image and a denoised one. Convolutional neural networks in combination with transfer learning can use the distinctive features of the residual noise to classify the image as real or fake. To get the resid-ual noise of an image, the image is denoised and then the denoised image is subtracted from the original [11], what is left is the noise of the original image.

[33] explore the PRNU-based image forensics methods in the context of deep-fake detection. The authors state that while a PRNU-based method is often outperformed by other deep learning methods, the method can be used in com-plement to other state-of-the-art methods. The basis for the method is the photo response non-uniformity (PRNU) sensor noise that is different for differ-ent types of devices [33]. [33] uses the PRNU's statistical properties to indicate content manipulation. The authors examine both spectral and spatial features of an image and use the energy, range, variance, skewness, kurtosis, variance in histogram, and position of the max value in the histogram. Which are then given to an SVM classifier which classifies a frame as real or fake. The classifier did reach a good accuracy on the tested datasets, however it cannot compete with state-of-the art deep learning methods [33]. In addition, the authors also examined the effects of cropping and extracted image and removing up to 50% of the image borders. This method improved accuracy on almost all tested datasets.

6.2.4. Generalizability

A problem a lot of deepfake detection methods have is being able to general-ize the model's ability to detect fake images from unknown source, i.e. real world samples. Currently, most detection methods are trained on a few known datasets from which the features are extracted, however this method does not account for unseen data. So when the trained model comes across a deepfake which does not fit the features of deepfakes in the training set, the model will be unable to recognize it. Aside from these methods focusing on specific artifacts,

[36] identify undirected approaches to detection as a valid method of detecting deepfakes. Some researchers treat the problem as a generic classification prob-lem and train deep neural networks as classifiers which decide the discriminatory features themselves. They train the classifier on both real and fake images, al-lowing the network to choose the features to discriminate upon. This method can lead to over fitting of the training set and might be unsuitable for use on real world data. The authors emphasize that this method of using classifiers for detection is flawed as adversaries can use adversarial machine learning methods to evade detection. A method that contrasts the classification method is the anomaly detection. These models are trained on normal data and then detect anomalies in the content on deployment. These methods therefore do not make any assumptions as to which attack could be used, making it more suitable as a general model to detect unknown creation methods [36].

[26] propose a method called OC-FakeDect to solve the generalizability issues of binary classification type deepfake detectors. Which are trained on datasets of real and fake images and require knowledge about the fake images to classify them. Detection performance on fake images from newer types of algorithms is therefore often mediocre [26]. propose to use a one-class Variational Autoen-coder (VAE) trained only on real face images, detecting fake images by finding anomalies and classifying them as fake. The images are evaluated through a met-ric called the anomaly score, which is the Root Mean Squared Error (RSME) between the input and output images of the VAE. By plotting the distribution of this score, the difference between real and fake images can be determined using statistical thresholding [26]. The authors test 2 different implementations of the same concept, the first implementation calculates the RSME directly be-tween the input and output images. While the second implementation uses an additional encoder and calculates the RSME between the latent representations of the images of both encodes. The authors say that by using the second method they hope to better capture the difference in characteristics between the images.

An example of how the second implementation can be seen in Fig. 3.

[45] propose using transfer learning on CNN architectures to make the detec-tion of deepfakes more generalized. By using transfer learning, a method which we will come back on later, the authors try to find a solution to the problem of overfitting in current detection methods. Transfer learning allows the model to leverage existing knowledge and data from a related domain to a new one [45], which should make the model function better on previously unseen data. In this case, a base model can be used in order to gather the first features available in a dataset, this knowledge can then be transferred to another model. This model then does not have to learn these features and can spend more time fine-tuning these neurons [45].

[4] propose to combine multiple CNNs for deepfake detection through a method called ensembling. Which is a machine learning method that enables models to be combined for a better prediction performance. The authors aim to train different CNN based classifiers to find different types of features in manipulated video frames that complement each other [4]. also implement an attention module which allows the user to see which part of the image was most important for the classification of that image. In addition, the authors ex-plore if Siamese training strategies can be used to infer more information about the data. This training strategy differs from most traditional training strate-gies which are end-to-end, which could prove useful in classification [4]. They describe the Siamese training paradigm as a method that aims to exploit the generalization capabilities of the networks to extract feature descriptions that emphasizes the similarities of samples in the same class. In other words, the fea-tures which are more descriptive of a certain class will be considered first during the classification, minimizing the time used for other, less discriminatory fea-tures. This differs from end-to-end training which extracts features end-to-end without considering whether a certain feature should be analyzed first.

[40] propose the Deepfake stack framework, which is an ensemble of deep learning methods to detect manipulated videos. The model combines multiple state-of-the-art detection methods into a single classification model [40]. create two layers which sit on top of each other, the first layer consists of pre-trained base learners and the second layer consists of a meta-learner which combines the learned features from the first layer and uses these to classify the video. The models are called level-0 models for the first layer and a level-1 model for the meta-learner. The meta learner learns through the predictions of the base layer on out of sample data, this is data that the models were not trained on [40]. This method can be used with any number of level-0 models, while a single meta leaner is needed. In this case the authors use XceptionNet, ResNet101, InceptionResNetV2, MobileNet, InceptionV3, DenseNet121 and DenseNet169 as base layers while the meta

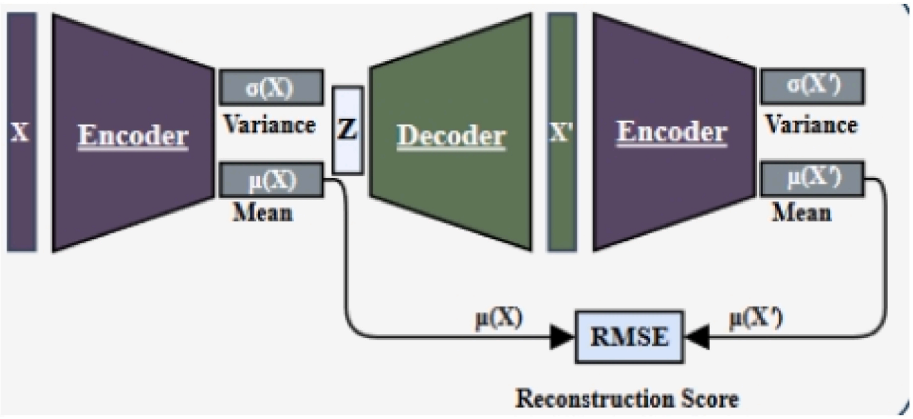

Fig. 3. OC-Fakedect-2 representation. Retrieved from [26].

learner is a CNN model called Deepfake Classifier (DFC). The ensembled model is called a metamodel, in which several sub-models provide a singular output. There are 2 types of ensembles, a stacking ensemble and a randomized weighted ensemble. The difference is in how the output of the sub-models are prioritized. In the randomized weighted ensemble, the models output is weighted based on the performance of the sub model of a validation dataset, while a stacking ensemble learns how to map the sub models predictions in such a way that it produces a better output than when randomly prioritizing the input [40]. A full representation of deepfakestack can be seen in Fig. 4.

[39] proposes an unsupervised approach to detecting GAN(deepfake) images. This approach works even without access to deepfake images during training, which makes it highly generalizable. The authors introduce several ideas in the paper that are the basis for the proposed system, called NoiseScope. One of the things mentioned is that similarly to camera fingerprinting, GAN model image generation leaves unique noise patterns related to the model that gener-ated the image. The model extracts this fingerprint to detect GAN images and is agnostic to the type of GAN used to generate the image [39]. An example of different types of fingerprints can be seen in Fig. 5. Although the picture might be obscure, their are very subtle differences in the pixels of each separate camera/GAN. These will return in any picture created with that camera or generated with that GAN. The proposed method should me more useful than supervised training methods as new GAN models are proposed very often, leav-ing supervised methods behind in detecting the new types of GAN images [39]. shows the difference between camera fingerprints, also known as photo-response non-uniformity (PRNU) patterns and images generated by GAN models.

[55] proposes a novel method of detecting deepfakes by capturing different levels of artifacts in a single network called a Multi-Layer Fusion Neural Net-work (MFNN). These levels are microcosmic or statistical features, mesoscopic features, and macroscopic features. Features from theses artifact levels are fused together before classification [55]. The authors note that the shallow layers cap-ture local anomalies in an image, deeper layers are more suited to capturing semantic anomalies and the middle layers are able to extract the mesoscopic features. These feature maps from the different levels of layers are sent directly to the last layer and concatenated to the feature vectors of the classification model [55]. By combining the features from all different layers, they have cre-ated a more generalized model that equally considers features from all layers.

6.2.5. *Long-Short-term-memory*

[18] is one of the first articles exploring using recurrent neural networks (RNN) to analyze both the frame-based features of a piece of content as well as the temporal features. The frame-based features are still learned through a Con-volutional neural network (CNN). Additionally, the authors use a Long-Short-term-memory (LSTM) network for temporal sequence analysis which is a type of recurrent neural network which unlike feedforward type neural networks, also utilizes feedback connections. It is therefore not only capable of image process-ing which consist of single data points, but also multiple data sequences such as video or speech data. It is very useful for temporal based detection algo-rithm. The CNN learns the frame-based features while the LSTM then learns the temporal features of a sequence of analyzed frames [18]. describe that while CNNs are successful in visual recognition tasks, LSTMs have long been used in sequence processing problems. The combination of a CNN and a LSTM within a deep learning architecture, is often considered to be both “deep-in-space” as well as “deep-in-time” [18].

[25] propose a method that uses a CNN and a C-LSTM network to detect deep-fakes. This is very similar to Ref. [18] which also built a model using the same methodology, however, the authors use a newer type of LSTM network for this proposed method. Nevertheless, the

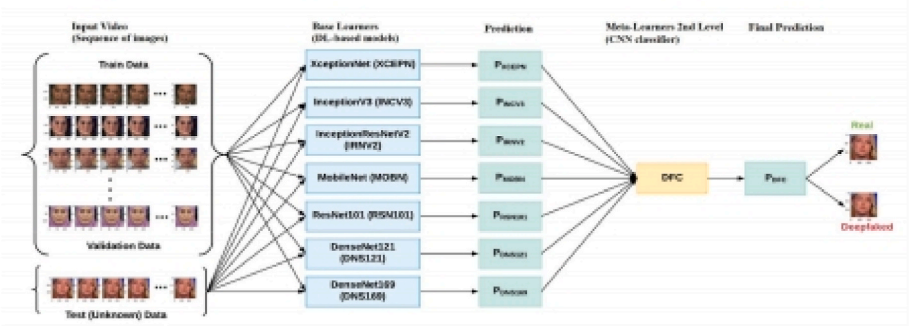

Fig. 4. Deepfakestack. Retrieved from [40].

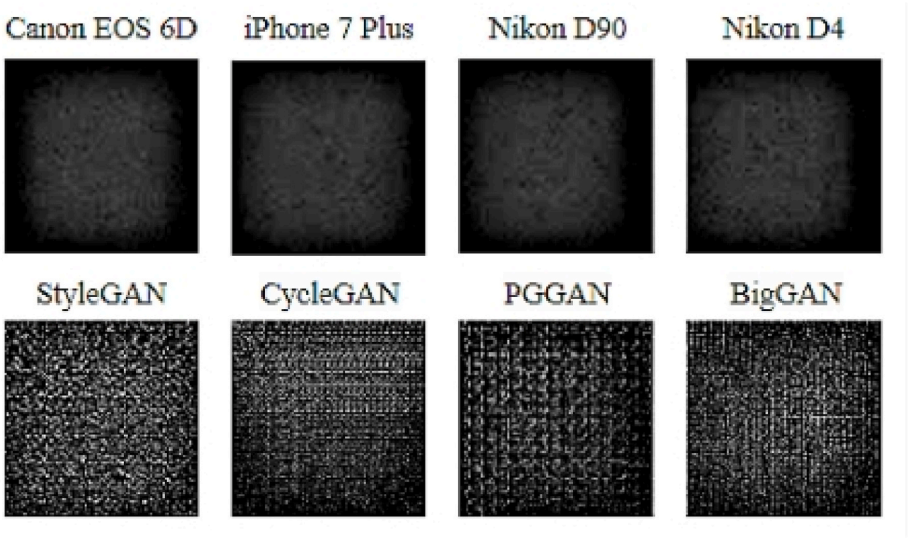

Fig. 5. Fingerprints examples. Retrieved from [39].

method is nearly the same [31]. proposes a deepfake detection method that uses 3d-attentional inception network to use both the spatial and temporal information with the 3d-kernels. In addition [31], applies an attention module to the CNN. The attention module is necessary for temporal based models because neural networks tend to forget information if the sequence of input data is too long. An attention module rectifies this problem by making the model consider all features.

[54] proposes a solution to deepfake detection by using a time-series perspec-tive to analyze two types of input features, Face-alignment (FA) and Dense- Face-Alignment (DFA). The temporal based methods both use a LSTM with an attention mechanism to analyze multiple frames [54]. The FA-LSTM utilizes 68 facial landmark points for analysis and has a fast inference speed, making it suitable for mobile applications. The DFA-LSTM uses pose adaptive features and measures changes in this. This method analyses the 3D space of the face to capture facial dense map changes, leading to a higher classification accuracy but lower inference speed due to higher complexity [54].

6.2.6. *Physiological signals*

[23] propose a novel way of detecting deepfakes through Eye-blinking patterns called Deepvision. Since blinking is an involuntary action based on a number of physiological and cognitive factors, significant changes or unnatural behavior in the blinking patterns can indicate that a video was manipulated. The au-thors used a General Adversarial Network and trained the model to recognize the biological factors (features) that determine Eye-blinking. The authors also provide the model with a plethora of metadata for pre-processing such as time-of-day, gender of the subject, age and activity that the subjects are engaged in. As these factors can significantly change the eye blinking patterns observed in the subjects [23]. detect the blinking between frame by using the eye-aspect- ratio(EAR) proposed by Ref. [43], which is a measure of the area taken up by the eyes. Its value reflects whether the eyes are opened or not, when the subject blinks, the EAR value of that frame suddenly drops and goes up again once the eyes are opened [23]. An example of this can be seen in Fig. 6. In this way, the blink frequency and time it takes to blink can be visualized in a graph and analyzed by the model [23]. designed the model to compare the blinking patterns in the pre-processed eye blinking database containing the metadata with the subject in the video and compares these based on activity level, time of day, age and gender.

6.2.7. *Spatio-temporal*

[32] propose a Channel-wise Spatiotemporal Aggregation (CSWA) module which uses consecutive video frames to fuse their deep features. This means that the features of multiple consecutive frames are fused together and evaluated with a shared weight. The module is used as a classifier on top of a CNN (EfficientNet B0) with skip connections to preserve low-level features which then extracts frame-level feature [32] s. The skip connections allows the model to remember the relevant frame level features which will also be used during classification. The CSWA module then fuses the feature maps obtained by the CNN and classifies the videos as either fake or real. Additionally, the authors find that cropping a slightly larger region (x1.3) around the face results in improved detection accuracy, since the model can differentiate better between pristine and manipulated pixels. This result remained consistent on all tested datasets.

[51] propose a novel method to detecting deepfakes based on analyzing video frames as a set to extract temporal inconsistencies as

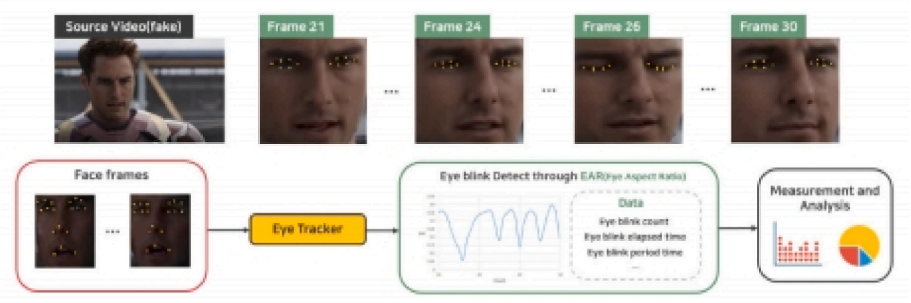

Fig. 6. EAR algorithm. Retrieved from [23].

well as extracting the in-dividual feature maps of all the frames. The authors introduce a new statistical based feature map operation called Set Reduce (SR), which aggregates features of an element in a set. The total pipeline uses multiple paths that the images take before reaching the final classifier. A set reduce path and multiple backbone paths with multiple changeable models, which use multiple feature extractors which will be used for classification in the discriminatory path [51].

In order to analyze the effects of Set Reduce on the total model, they use varying numbers of set reduce operations to investigate the best number to use. The conclusion of this experiment is that it varies with the underlying models that are used. Some models perform better with fewer set reduce operations while some models fare better with more [51]. However, all of the models per-formed better with some form of Set reduce operations in contrast to none at all, suggesting that the method proposed by Ref. [51] is of value to future detection efforts.

[3] proposes a new method for detecting deepfakes based on the optical flow of multiple frames. An example of which can be seen in Fig. 7. They extract optical flow fields containing inter-frame correlations which a CNN classifier can use as input for classification. Specifically, optical flow describes using a vector matrix field on two consecutive frames and comparing the motion across frames. The author hypothesized that synthetically created motions would differ from those naturally captured by a camera [3]. feed two matrices into a semi trainable CNN called flow-CNN. The results from initial experimentation presented by the authors shows that the method has promising results and could be further developed in a fully-fledged deepfake detection method. The method proposed by Ref. [3] is different from other spatio-temporal based methods in that you only need 2 frames to classify a video. Whereas other spatio-temporal methods often need a lot of consecutive frames to classify the video correctly. This method may therefore be a lot more efficient compared to non-frame based methods.

[46] propose a novel deepfake detection method based on spatial and tempo-ral characteristics of a video to be better generalizable to Deepfake-in-the-Wild (DFW) videos. The authors note that common deep learning methods for de-tecting Deepfakes lack real-world transferability and perform poorly on DFW videos [46]. proposed an architecture based on convolutional LSTM cells com-bined with Residual blocks. Which is capable of utilizing both the spatial & temporal information available in consecutive frames. In addition, the authors propose a new training strategy aimed to improve the generalizability of deep-fake detection methods and included a new DFW dataset that can be used to benchmark deepfake detection methods on DFW videos [46]. outline three different threat models that are relevant to modern day deepfake detection. These threat models can be seen in Fig. 8. In the first threat model, the in-domain-attack, the dataset that the attacker used to create the Deepfake is the same as the training set used in the detector. In a real world scenario, this is highly unlikely. In the second threat model, the out-of-domain-attack, the dataset that the attacker used is a known dataset, but different from the one used to train the detector. This is a far more likely scenario. In the last threat model, the open-domain-attack, the deepfake generated by the attacker is based on unknown datasets, i.e. DFWs [46]. While the detector is trained on a known dataset. This is the most likely real-world scenario, while most deepfake de-tection schemes are unequipped to deal with such attacks. The authors then outline 3 different training strategies and how they relate to the threat models. Single domain learning is training the detector on a single dataset, while Merge learning is training the detector on multiple known datasets, which can decrease the accuracy on detection of these deepfakes as the model has to learn to many different features. Both of these types of training are insufficient for the defence against open-domain attacks. Therefore the authors propose a type of transfer learning, where initially the model is only trained on a single dataset like in sin-gle domain learning, and is subsequently trained on very small samples of other known datasets, only learning the anomalies of these other datasets compared to the original training set.

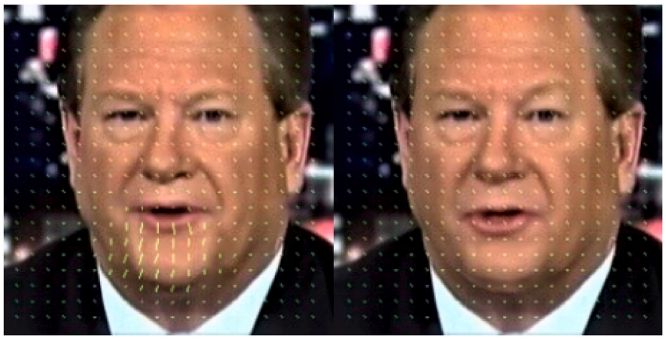

Fig. 7. Optical flow difference between normal(left) and Deepfake(right). Retrieved from [3].

[46] confirm this theory during experimentation, the transfer learning method had considerably higher accuracy's across multiple different datasets than the single domain learning method as well as the merge learning methods.

6.2.8. *Vanishing gradient problem*

[20] describes the vanishing gradient problem, which is the problem that when back-propagating through a neural network the partial derivatives of each layer are calculated, when the value of the partial derivative starts getting lesser then one, the weights of that layer of neurons is not properly adjusted [20]. This value gets lower as the back propagation gets closer to the beginning, which means that the value of the weights of the features examined at the beginning of the network are not properly adjusted and thus not taken into account in the final model [20].

6.2.9. *Platforms*

[30] proposes an open web-based platform called Deep-fake-o-meter which hosts more than 10 state-of-the-art detection models to help anyone with detecting deepfakes. It can be used to compare the efficiency of multiple Deepfake de-tection algorithms on a single input. The authors provide the design of an architecture of the front-end, back-end and data synchronization components. A user can upload a video, select the desired detection methods, and input his email. After processing is done on the back-end of the software, the user will re-ceive an email containing detailed results regarding detection [30]. This method is useful for developers of deepfake detection algorithms to run their algorithms on a remote server. As well as researchers looking to benchmark their own de-tection algorithms against state-of-the-art methods and users to detect whether a certain video is a deepfake or not.

[42] propose a deepfake detection system aimed at journalists. The authors choose a video detection model and construct their own temporal based detec-tion model including frame-level artifacts, inspired by an RNN-based deepfake detection model. In addition, they explore the capabilities of fake audio de-tection model. Both models showed a good AUC for fake video and audio detection, however the paper was aimed at journalists, so the main contribution of the paper was the design of an intuitive application which can be used by journalists to determine whether a video is real or fake.

6.3. *Improving detection methods*

In order to improve the defenses against Deepfakes in digital forensics, re-searchers sometimes take on the role of the attacker to expose vulnerabilities of current methods. This often leads to increased efforts to cover these vulnerabil-ities and improves the overall defense against deepfakes.

[7] expose multiple vulnerabilities of modern day deepfake detection methods by implementing two types of attacks specifically to confuse neural network classifiers. The white-box attack can reduce the classification accuracy of a specific classifier to almost 0% [7]. In addition, the

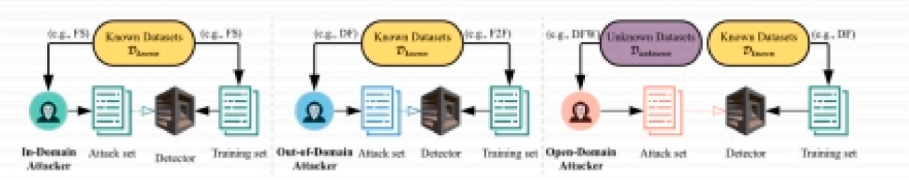

Fig. 8. Different domains. Retrieved from [46].

authors develop a black-box attack that reduces a random classifiers area under the Curve (AUC) to below 0.2, showing that most state-of-the-art detection methods are vulnerable to specific attacks. The authors state that they do not wish to weaken the detection efforts but that they are mainly concerned with strengthening the current forensic methods, therefore they did not release the source code used for the attacks [7]. use a white-box method called Distortion-minimizing attack to attack a well-known detection method from Ref. [48] by modifying some pixels automatically that are used to minimize the loss function in the classifier. Automatically targeting these specific pixels makes it so that you only have to change a small percentage(4%–11%) of pixels to confuse the classifier and make it misclassify 89.7%–100% of all fake images [7]. The optimal number of pixels is calculated using a formula that optimizes the amount of fake images misclassified while minimizing the amount of modified pixels. In addition, two other white-box attack methods tested on different detection methods are described by the authors, which perform similarly to the method described above. The Black-box transfer attack is more relevant to real world scenario's since in the real world the adversary would not know what kind of detection method will be used to classify the deepfake. They describe that in this case the adversary will not be able to simply use gradient-descent optimizations to find the optimal attack [7]. assume that in the black-box scenario, the adversary is aware that some type of defense is being used and that he knows the general strategy of the defense. The authors built their own classifier and created an attack based on their own classifier, then transferred this method to detection method of [48] which reaches a significantly higher accuracy than the authors' method. The transferred attack still reduced the AUC of the classifier from 0.96 to 0.22 [7]. Which is weaker than the white-box attacks, but still very significant in a real world scenario, as this makes the classifier completely unreliable. An example both of these types of attacks can be seen in Fig. 9.

6.4. *Datasets*

The deepfake detection research community has brought forth a lot of datasets which can be used as training and testing sets as well as provide a benchmark against other deepfake detection methods. This section will provide an overview of the often used and most recent datasets.

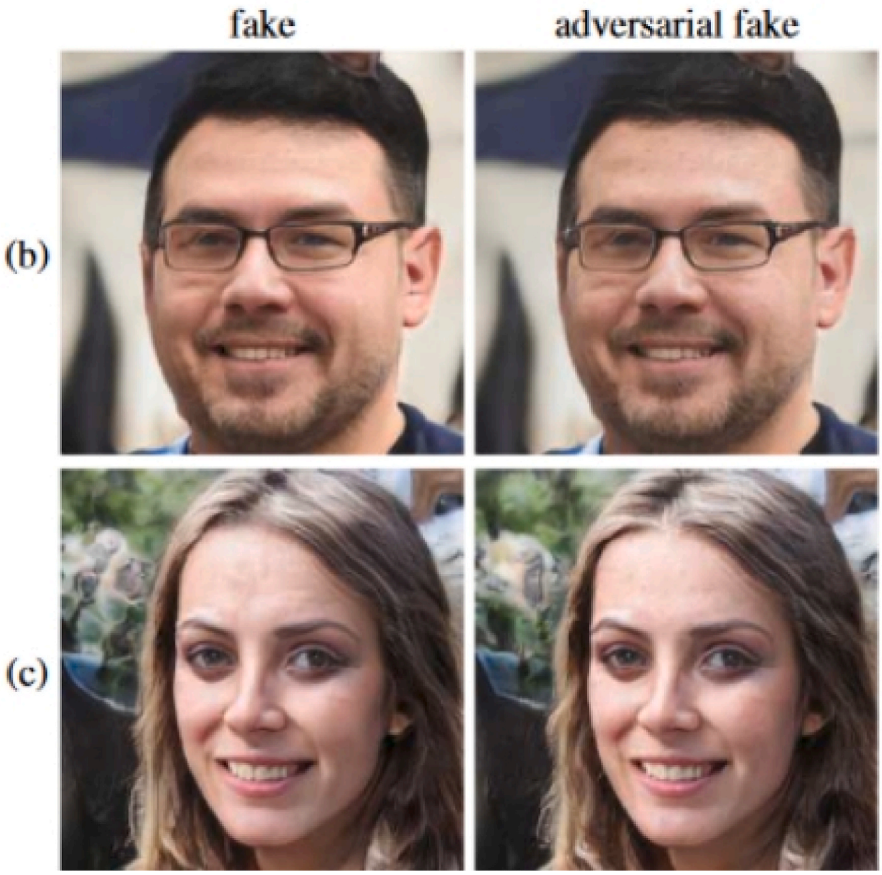

Fig. 9. Adverserial fakes. Retrieved from [7].

6.4.1. FaceForensics++

The most commonly used dataset is the FaceForensics++ (FF++) dataset from Ref. [41], which combines several other datasets into a single dataset [41]. created an automated benchmark for facial manipulation detection. The benchmark includes Deepfakes, Face2Face, FaceSwap and NeuralTextures as the facial ma-nipulation standard at random levels of compression and size. This research is relevant to the deepfake detection efforts as detection algorithms can now be benchmarked against each other and real-world data. The benchmark is also measured against a human baseline [41].

[41] shows that they can reliably detect facial manipulations and outperform humans by using convolutional neural networks to extract image features. The neural network in particular is trained in a supervised way, which means that a large training and testing set of 1.8 m images extracted from a 1000 videos is created with ground truth classifications [41]. The large scale dataset of facial manipulations is created based on methods like Face2Face and Faceswap and also including Deepfakes and Neuraltextures. The generated dataset is then compressed using the most common method of compression in social media applications, to create high-quality and low quality videos [41].

The detection methods are benchmarked against a human test were it was treated as a binary classification problem [41]. For the most part, the humans test barely outperformed random chance for most types of manipulations, even performing worse than random chance in the NeuralTextures accuracy. To benchmark the detection methods [41], manually selected one challenging frame from each of the 1000 videos. Either from a manipulated video or the original footage. Most detection methods explored in the article seem to perform signifi-cantly worse on lower quality videos. Reaching relatively high accuracy's on the pristine and high quality compressed videos, but much lower accuracy on the low quality compressed videos. This presents a challenge to detection methods to improve the detection of low quality videos were artifacts may not be very obvious.

6.4.2. CelebDF

[29] present a new large-scale high quality deepfake dataset called CelebDF, containing 5639 challenging deepfake videos of celebrities generated through an improved synthesis process. This dataset is widely used by many modern state-of-the-art detection methods to assess the accuracy and efficiency of the detection method compared to other detection methods. The aim of the authors was to bridge the gap between the quality of deepfakes in detection datasets and actual deepfake videos found on the internet. The datasets used at the time of writing the article had lots of obvious artifacts and weirdly synthesized faces which can easily be identified by humans. So high detection accuracy on these datasets may not be relevant when these methods are actually used on real deepfakes [29]. also include 590 real videos which are from youtube videos containing celebrity interviews. The Deepfakes are made using an improved Deepfake creation algorithm. With several measures in place to reduce tradi-tional problems with deepfakes such as a color mismatch, temporal flickering and inaccurate face mask [29]). An example of images from CelebDF can be seen in Fig. 10. The images on the left are genuine, all other images are deep-fakes.

6.4.3. Deepfake detection challenge

In 2018 AWS, Facebook, the Partnership on AI's Media Integrity Steering Com-mittee, and academics all came together to create the Deepfake Detection chal-lenge (DFDC) [24]. With one millions dollars as prize money, the challenge was aimed at increasing effort spent into deepfake detection research. Along with the challenge, a new dataset was released consisting of over 470gb of training data [24]. The dataset was created with paid actors and features over 100.000 clips from several different deepfake creation algorithms [24]. It is currently the largest available dataset for deepfake detection.

6.4.4. Other datasets

In addition to the datasets mentioned earlier, there are several other recent and often used datasets such as Face2Face (F2F), DeepfakeTIMIT, FaceSwap (FS), DeepfakeDetection (DFD). Some of these datasets are a bit older and are not used that often anymore.

Fig. 10. CelebDF example from [29].

6.5. Prevention

In addition to deepfake detection methods, digital forensics research also con-cerns itself with alternative methods to ensure media authenticity.

6.5.1. Proof of authenticity

Proof of authenticity is another theme that emerged in the literature. Instead of detecting fake videos, proof of authenticity systems try to verify that a piece of content is real at the moment it is made [13]. explore the potential of using blockchain and distributed ledger technologies (DLT) to guarantee authenticity and traceability in the fight against digital deception. In the realm of foren-sic methods to combat deepfakes, this method is a preventative method. The authors' research does not limit itself to just deepfakes and instead dives into all forms of online misinformation [13]. examine a smart contract based au-thenticity verification system which provides a number of capabilities to combat digital deception. The DLT-based system would provide a method for decen-tralized content moderation, contrastingly to conventional content moderation, this would allow a large group of validators to provide consensus over the au-thenticity of a piece of digital content [13]. Another capability provided by the system is a trustworthiness checker, a system where a piece of content can be verified as part of the blockchain by any node in the network. The sys-tem could also incentivize fact-checkers by providing financial rewards as well as reputation for high-quality fact checking. This reputation system could also be used to measure the credibility of a certain publisher. Readers can see the reputation of certain publishers and be warned when their reputation is low due to publishing a lot of fake news. The system could also provide a place for decentralized social media platforms which would improve privacy as well as data ownership, making it harder for misinformation about certain persons to spread [13]. An overview of the DLT platform and it's capabilities can be seen in Fig. 11.

In addition to the aforementioned benefits of the system, DLT could also provide a home for a variety of services related to data provenance and *trans*-parency. It could demonstrate the origin of content and detect counterfeit con-tent, forging content would be almost impossible in this scenario [13]. News would also be traceable to the original publisher using the meta-data, times-tamps and the links between different blocks, allowing users to trace the news back to its source and verify it's integrity. They also mention a few challenges and drawbacks of the proposed DLT system. Current research is mostly focused.

Around verifiable fake content, while other types of fake news that cannot be proven false are ignored. Scalability of such a system may also prove to be a major problem, as the level of required decentralization needed to achieve the benefits of the system could impact its performance. Additionally, current pro-posals are mostly based on cryptographic hashes, which completely change if for example a single pixel of an image is changed, resulting in reduced traceability of content [13]. This problem is especially relevant to deepfakes which swap out multiple pixels in a piece of content. The problem can be resolved by using dif-ferent methodology for storing content on the blockchain, but current proposed methods are not capable of these alternative methods. In conclusion, the use of DLT's would make handling of disinformation, including deepfakes, a lot easier by providing a variety of methods for authenticity, verification, transparency and traceability [13].

[8] proposes a proof of authenticity (PoA) method based on blockchain tech-nology to prevent deepfakes from spreading misinformation. The proposed framework would require multimedia content to be hashed and deep encoded with CNN-LSTM as a PoA blockchain based service to ensure its unaltered authenticity. The authors employ IBM's Hyperledger Fabric framework 2.0 so that users still have sovereignty over their media content. On this blockchain, a user would be able to trace the Media back to its source and verify its authen-ticity by viewing the changes a piece of content has undergone. This method lacked scalability, since maximum transaction payload was limited to 100 mb as the Hyperledger Framework was only able to handle these transaction loads [8].

However, it may prove useful for highly sensitive content such as presidential speeches, which on a distributed ledger with a representative hash would be able to be authenticated. This method would also allow owners of content to grant rights to specific users to use and modify their content, through a smart contract-based system [8].

6.5.2. Active defense

In [44] the problem of deepfakes is approached in a different way than detec-tion. The authors aim to disrupt facial landmark extraction of the input images of deepfake generation software so that the deepfakes that are created are of lesser quality and easier to detect. They achieve this by using adversarial per-turbations of the image. A method which they call landmark breaker, since it is supposed to break the state-of-the-art facial landmark extractors. This method could protect

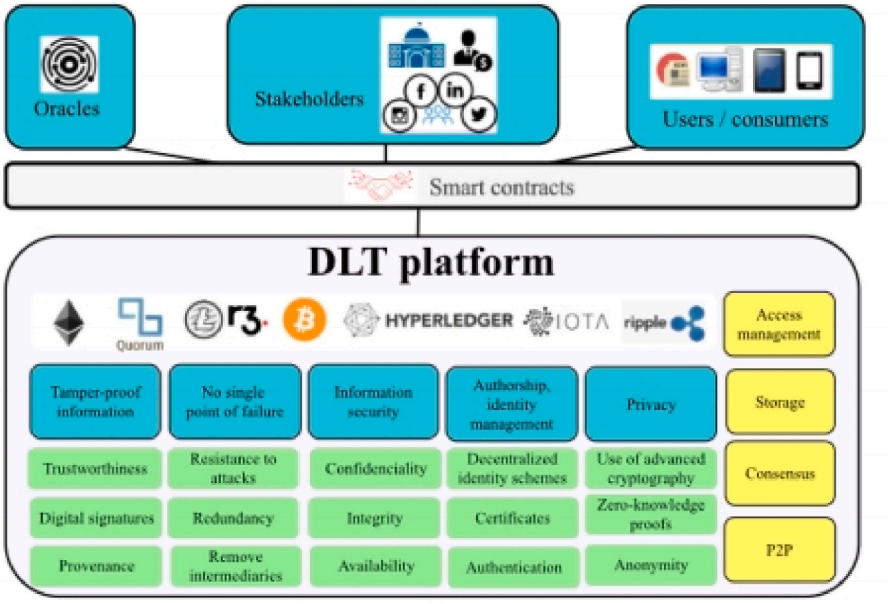

Fig. 11. DLT platform from [13].

videos from being manipulated as the landmarks which the network uses to synthesize the faces on are detected incorrectly, which leads to a weird, obviously fake result [44]. The method makes use of a new loss function which aims to make the error between predicted and original heatmaps more likely [44]. Initial experiment performed by the authors on the Celeb-DF dataset with multiple state of the art facial landmark extractors proof that the method is effective in disrupting landmark extraction. The resulting deepfakes from these algorithms are obviously fake. Similarly to Refs. [34,44] proposes a watermark-based system to prevent images from being manipulated by a neural network. The watermarked image turns into a blurry image when it is manipu-lated by a neural network, as the features of the image are extracted incorrectly [34]. The watermark is not visible for humans in the original image but will disrupt the manipulation of the image anyway. The watermark module used in this method is built with a convolutional network which extracts the facial information and a deconvolutional network generates the watermark of the im-age [34]. An attention module is used to only apply the watermark in the facial area of the image. This is not the first watermark-based model, but in the results, this model is significantly faster for the Celeb-DF dataset than other state-of-the-art watermarking methods.

6.5.3. Social verification

[47] propose a social verification method to combat deepfakes in which video's of speakers are captured from multiple different angles. The facial geometry consistency is then analyzed and potential manip-ulations will be easily detected. The authors show that they can detect small manipulations such as subtle mouth movements out of a set of 6 videos using 2d mouth landmark motion. This setup might be useful in the battle against live deepfakes in which the speakers face is being deepfaked in real time. As the use of 6 different viewing angles will make this hard to achieve in addition to the manipulation detection method. In addition to the method, the author also created their own dataset consisting of 25 speakers from multiple viewing angles, with subtle manipulations in some of the videos.

6.6. Conclusion

There are many methods currently available to forensic analysts to prevent the negative effects of deepfakes. Current deepfake detection methods seem to be an effective tool for detecting fake videos. The machine learning models can reach high detection accuracies on large datasets of high-quality deepfakes. However, most state-of-the-art detection methods are not equipped to handle out-of-domain data and deepfakes in the wild. Which is supposed to be the main use of deepfake detection algorithms. Currently, the detection algorithms, save a few specifically designed for DFWs, are hampered by the way they are trained and the public datasets that are used to validate these detection methods. It is a difficult problem to handle out-of-domain data, although some detection methods such as [26,46] are better equipped for this type of detection. GANs allows for the easy creation of relatively high quality deepfakes. The re-search into detection techniques also fuels the development of better technology which is able to avoid detection. This arms race in detection and creation has enabled high quality, hard-to-detect deepfakes to emerge. These are not easily created and have to be made by someone with experience and expertise on the subject, but the development of the GANs will allow for high-quality deepfakes to be made by anyone. The existence of these high-quality deepfakes which might be hard to detect means that not only will videos have to be proven fake, it will also be necessary to be able to prove that videos are real. Which cur-rent most of the current detection methods are unable to do due to the lack of transparency in how the model got its results. Current AI methods often don't offer an explana-tion as to what clues it used to determine why a frame is con-sidered fake. Deepfake detection will always be behind the deepfake creation, as the detection techniques are reactionary by definition. Anyone with mali-cious intent could create a very realistic deepfake that easily evade detection in modern state of the art methods [7]. has also shown that current detection methods are still very vulnerable to specific attacks such as adverserial fakes. Solving this problem will prove difficult if we keep training and benchmark-ing the detection methods in the same way. Generalizability should therefore be the most important research topic. Blockchain and smart contract methods may be a solution, although current implementations of this are not scalable and severely lacking in the tools needed to execute this type of proof-of-authenticity platform. In addition, the required amount of adoption of such a system to make it work will not be realistically reached anytime soon. In conclusion, the current available methods to prevent the negative effects of deepfakes are not sufficient. It is still hard to prove why a video is real, which is relevant in a forensic scenario. Authenticating content when it is created might be a solution in the future. However, current imple-mentations are not scalable, will require widespread adoption, and offer no solutions to deepfakes in the short term.

7. Gait analysis

The field of forensic gait analysis is researched in United Kingdom, Australia, India China and the Netherlands. Also in this field over 7000 hits are found with google scholar from 2018. For this reason a selection of the most relevant papers is given.

In the United Kingdom standardization is sought for gait analysis [127] were a code of practice is written. Professor Birch in the United Kingdom has published many papers and even a book on the accuracy of gait analysis, since it appears to be overstated by some others ([63–67, 91]).

In the Netherlands, the University of Groningen has been active on modeling gait also with a modeling tool for gait as well as measuring it with a model [143]. Also, the Netherlands Forensic Institute has written a critical review article on the state of the art in 2018 [141].

In Australia, there were also issues with gait analysis in court ac-cording to Ref. [80]. If presented incorrectly in court it can provide miscarriages of justice. The evidence from gait analysis should also not be overstated according to research from Seckiner [133,134]) and also CCTV distortions and artifacts have to be taken into account.

For research databases [139] have been provided by different groups (Birch et al., 2018; Makihara et al., 2020; Uddin et al., 2018). Also deep learning is used for training the algorithms based on these databases [88, 115,126,130,131]. Click or tap here to enter text. The practical use in forensic science appears to be limited currently of these trained networks.

The history of gait analysis as evidence is described by Nirenberg [122,123] Often gait analysis is used as last resort, if face coverings are used [138].

The University of Copenhagen also did a retrospective review on their research on gait analysis, since they were one of the first publishing on this in forensic literature [121].

The standard of evidence admissibility in the United States also has been discussed [104,105] Furthermore, there are many papers that discuss the reliability of gait analysis [60,68,72,89,93,94,96,105,106, 108,116,120,123,125,126,129,130,132,136,140,144].

8. Conclusion

Gait analysis in CCTV can be conducted, however quality assurance and adhering to standards is important. Also if multiple cameras are available this can help. Gait is a variable biometric, Since people wear clothes most often, some of the features that can be used in gait com-parison are concealed, so attention should be given to these.

9. ENF in Video

The Electric Network Frequency (ENF) in Video is available by

several methods. It can either be present in the audio signal or in the video signal itself by the rolling shutter effect [71,84]. The ENF is based on the variation in the 50 or 60 Hz on the mains, and by comparing it with the recordings of the variations an estimation can be provided when a recording has been made.

Some validation based on different light sources is described by Frijters and Geradts [81]. With high speed cameras the estimation is better, since with higher frame rates more variations in the ENF [77,78].

For practical use of ENF error analysis is important according to Hua [87], since one needs enough data to draw conclusions. Also ENF for intra-grid location estimation is discussed in several papers [82, 117–119].

The method can be used for forging attacks [118] and detecting malicious frame injection attacks in CCTV [119].

Other authors report the use in smart cities [146] Though also the recording device identification is described [69].

10. Conclusion

ENF in video by the use of the rolling shutter effect can in optimal conditions be used for estimation of when a recording has been made. Validation remains important and care has to be given not to overstate conclusions.

11. Photo Response Non Uniformity (PRNU)

The use of Photo Response Non Uniformity is well known to work on camera identification (same source or different source) and also determining with which camera model a video has been captured.

The method has been used also in deepfake research in several instances, which provides an additional way of research [103,142,148, 149].Click or tap here to enter text.

There is also some criticism on the PRNU pattern itself and its uniqueness as such. [73] [98,99,101].

Also newer video databases are developed, which are important for the field [79,83]. [85,135] since the pattern changes with new cameras, as also can be seen with the newer smartphones.

Knowledge of the video format is important to know which frames should be selected for the PRNU [59,77,97,100,101,111].

Recompression by social media platforms might alter the PRNU pattern itself [95,107]Click or tap here to enter text.

Publications on using AI in camera model determination from video forensics is attracting attention [57,58,92,113,145].

Finally new methods for PRNU extraction that are faster or more useful for certain settings (such as motion compensation) in video cameras are important since the rate of change in this field is fast [74, 101,110,111,114,147].

12. Conclusion

The field of PRNU in video forensics is moving fast. It can be used for manipulation detection as well as camera source determination. Due to motion compensation methods in the cameras as well as compression, faster and better algorithms are needed for PRNU, and much research is available. There is some criticism on the uniqueness of the PRNU pattern, and for this reason validation by using same camera and same model references is advised.

References

[1] D. Afchar, et al., MesoNet: a compact facial video forgery detection net- work, WIFS (2018), https://doi.org/10.1109/WIFS.2018.8630761, 2019, 10th IEEE International Workshop on Information Forensics and Security.

[2] Atmik Ajoy, et al., DeepFake Detection using a frame based approach involving CNN, in: 2021 Third International Conference on Inventive Research in Computing Applications, ICIRCA), 2021, pp. 1329–1333, https://doi.org/10.1109/ICIRCA51532.2021.9544734.

[3] I. Amerini, et al., Deepfake video detection through optical flow based CNN, ICCVW (2019) 1205–1207, https://doi.org/10.1109/ICCVW.2019.00152, 2019, Proceedings - 2019 International Conference on Computer Vi- Sion Workshop.

[4] Nicolò Bonettini, et al., Video face manipulation detection through En- semble of CNNs, in: 2020 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 5012–5019, https://doi.org/10.1109/ICPR48806.2021.9412711.

[5] Jason Brownlee, A Gentle Introduction to Generative Adversarial Net- Works (GANs), 2019 url: https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/.

[6] R. Caldelli, et al., Optical Flow based CNN for detection of unlearnt deep- fake manipulations, in: Pattern Recognition Letters 146, 2021, pp. 31–37, https://doi.org/10.1016/j.patrec.2021.03.005.

[7] N. Carlini, H. Farid, Evading deepfake-image detectors with white- and black-box attacks, in: IEEE Computer Society Conference on Com- Puter Vision and Pattern Recognition Workshops, 2020, pp. 2804–2813, https://doi.org/10.1109/CVPRW50498.2020.00337, 2020-June.

[8] Christopher Chun, Ki Chan, et al., Combating deepfakes: multi-LSTM and blockchain as proof of authenticity for digital media, in: 2020 IEEE/ITU International Conference on Artificial Intelligence for Good (AI4G), 2020, pp. 55–62, https://doi.org/10.1109/AI4G50087.2020.9311067.

[9] Polychronis Charitidis, et al., Investigating the Impact of Preprocessing and Prediction Aggregation on the DeepFake Detection Task", 2020.

[10] A. Chintha, et al., Recurrent convolutional structures for audio spoof and video deepfake detection, in: IEEE Journal on Selected Topics in Signal Processing 14, 2020, pp. 1024–1037, https://doi.org/10.1109/JSTSP.2020.2999185, 5.

[11] M.C. El Rai, et al., Fighting deepfake by residual noise using convolu- tional neural networks, in: 2020 3rd International Conference on Signal Processing and Information Security, ICSPIS 2020, 2020, https://doi.org/10.1109/ICSPIS51252.2020.9340138.

[12] S. Ferreira, M. Antunes, M.E. Correia, Exposing manipulated photos and videos in digital forensics analysis, J. Imag. 7 (2021) 7, https://doi.org/10.3390/jimaging7070102.

[13] P. Fraga-Lamas, T.M. Fernandez-Carames, Fake news, Disinforma- tion, and deepfakes: leveraging distributed ledger technologies and blockchain to combat digital deception and counterfeit reality, in: IT Professional 22.2, 2020, pp. 53–59, https://doi.org/10.1109/MITP.2020.2977589.

[14] R.A. Frick, S. Zmudzinski, M. Steinebach, Detecting deepfakes with Haralick's texture properties, in: IS and T International Symposium on Electronic Imaging Science and Technology 4, 2021, https://doi.org/10.2352/ISSN.2470-1173.2021.4.MWSF-271.

[15] R.A. Frick, S. Zmudzinski, M. Steinebach, Paper: detecting "Deep- Fakes" in H.264 video data using compression ghost artifacts", in: IS and T International Symposium on Electronic Imaging Science and Technol- Ogy 4, 2020, https://doi.org/10.2352/ISSN.2470-1173.2020.4.MWSF-116, 2020.

[16] Ian Goodfellow, et al., Generative adversarial nets, in: Advances in Neural Information Processing Systems 27, 2014.

[17] L. Guarnera, et al., Preliminary forensics analysis of DeepFake images, AEIT (2020), https://doi.org/10.23919/AEIT50178.2020.9241108, 2020, 12th AEIT International Annual Conference.

[18] D. Guera, E.J. Delp, Deepfake video detection using recurrent Neu- ral networks, in: Proceedings of AVSS 2018 - 2018 15th IEEE Interna- Tional Conference on Advanced Video and Signal-Based Surveillance, 2019, https://doi.org/10.1109/AVSS.2018.8639163.

[19] Z. Guo, et al., Fake face detection via adaptive manipulation traces ex- traction network, in: Computer Vision and Image Understanding 204, 2021, https://doi.org/10.1016/j.cviu.2021.103170.

[20] Yuhuang Hu, et al., Overcoming the Vanishing Gradient Problem in Plain Recurrent Networks, 2018 *arXiv preprint arXiv:180*1.06105.

[21] A. Ismail, et al., A new deep learning-based methodology for video deep- fake detection using xgboost, Sensors 21 (2021) 16, https://doi.org/10.3390/s21165413.

[22] M.T. Jafar, et al., Digital forensics and analysis of deepfake videos, in: 2020 11th International Conference on Information and Communication Systems, ICICS 2020, 2020, pp. 53–58, https://doi.org/10.1109/ICICS49469.2020.239493.

[23] T. Jung, S. Kim, K. Kim DeepVision, Deepfakes detection using human eye blinking pattern, in: IEEE Access 8, 2020, pp. 83144–83154, https://doi.org/10.1109/ACCESS.2020.2988660.

[24] Kaggle Deepfake, Detection Challenge, 2018 url: https://www.kaggle.com/c/deepfake-detection-challenge/overview (visited on 11/17/2021.

[25] S. Kaur, P. Kumar, P. Kumaraguru Deepfakes, Temporal sequential analysis to detect face-swapped video clips using convolutional long short- term memory, J. Electron. Imag. 29 (2020) 3, https://doi.org/10.1117/1.JEI.29.3.033013.

[26] H. Khalid, S.S. Woo, OC-FakeDect: classifying deepfakes using one- class variational autoencoder, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 2794–2803, https://doi.org/10.1109/CVPRW50498.2020.00336.

[27] E. Kim, S. Cho, Exposing fake faces through deep neural networks combining content and trace feature extractors, in: IEEE Access 9, 2021, pp. 123493–123503, https://doi.org/10.1109/ACCESS.2021.3110859.

[28] X. Li, et al., Sharp multiple instance learning for DeepFake video de- tection, in: MM 2020 - Proceedings of the 28th ACM International Con- Ference on Multimedia, 2020, pp. 1864–1872, https://doi.org/10.1145/3394171.3414034.

[29] Y. Li, et al., Celeb-DF: a large-scale challenging dataset for DeepFake forensics, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020, pp. 3204–3213, https://doi.org/10.1109/CVPR42600.2020.00327.

[30] Y. Li, et al., DeepFake-o-meter: an open platform for DeepFake detec- tion, in: Proceedings - 2021 IEEE Symposium on Security and Privacy Workshops, SPW, 2021, pp. 277–281, https://doi.org/10.1109/SPW53761.2021.00047, 2021.
[31] Changlei Lu, et al., Deepfake video detection using 3D-attentional Incep- tion convolutional neural network, in: 2021 IEEE International Con- Ference on Image Processing, (ICIP), 2021, pp. 3572–3576, https://doi.org/10.1109/ICIP42928.2021.9506381.
[32] Y. Lu, et al., Channel-wise spatiotemporal aggregation technology for face video forensics, in: *Security And Communication Networks* 2021, 2021, https://doi.org/10.1155/2021/5524930.
[33] F. Lugstein, et al., PRNU-Based deepfake detection, in: IH and MMSec 2021 - Proceedings of the 2021 ACM Workshop on Information Hiding and Multimedia Security, 2021, pp. 7–12, https://doi.org/10.1145/3437880.3460400.
[34] L. Lv, Smart watermark to defend against deepfake image manipu- lation, in: 2021 IEEE 6th International Conference on Computer and Communication Systems, ICCCS 2021, 2021, pp. 380–384, https://doi.org/10.1109/ICCCS52626.2021.9449287.
[35] M. Masood, et al., Classification of deepfake videos using pre-trained convolutional neural networks, in: 2021 International Conference on Digital Futures and Transformative Technologies, ICoDT2 2021, 2021, https://doi.org/10.1109/ICoDT252288.2021.9441519.
[36] Yisroel Mirsky, Wenke Lee, The creation and detection of deepfakes: a survey, in: ACM Computing Surveys (CSUR) 54, 2021, pp. 1–41, 1.
[37] B.F. Nasar, T. Sajini, E.R. Lason, Deepfake detection in media files - audios, images and videos, in: 2020 IEEE Recent Advances in Intelligent Computational Systems, RAICS 2020, 2020, pp. 74–79, https://doi.org/10.1109/RAICS51191.2020.9332516.
[38] N.F.I. Vakbijlage, Waarschijnlijkheidstermen, url: https://www.forensischinstituut.nl/over-het-nfi/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen, 2020.
[39] J. Pu, et al., NoiseScope: detecting deepfake images in a blind setting, in: ACM International Conference Proceeding Series, 2020, pp. 913–927, https://doi.org/10.1145/3427228.3427285.
[40] M.S. Rana, A.H. Sung DeepfakeStack, A deep ensemble-based learn- ing technique for deepfake detection, in: Proceedings - 2020 7th IEEE International Conference on Cyber Security and Cloud Computing and 2020 6th IEEE International Conference on Edge Computing and Scalable Cloud, CSCloud-EdgeCom 2020, 2020, pp. 70–75, https://doi.org/10.1109/CSCloud-EdgeCom49738.2020.00021.
[41] Andreas Rössler, et al., FaceForensics++: Learning to Detect Manipulated Facial Images, 2019 arXiv: 1901.08971[cs.CV].
[42] S.J. Sohrawardi, et al., Poster: towards robust open-world detection of deepfakes, in: Proceedings of the ACM Conference on Computer and Communications Security, 2019, pp. 2613–2615, https://doi.org/10.1145/3319535.3363269.
[43] Tereza Soukupova, Cech Jan, Eye blink detection using facial land- marks, in: 21st Computer Vision Winter Workshop, Rimske Toplice, 2016. Slove- nia.
[44] Pu Sun, et al., Landmark breaker: obstructing DeepFake by disturbing landmark extraction, in: 2020 IEEE International Workshop on Infor- Mation Forensics and Security, WIFS), 2020, pp. 1–6, https://doi.org/10.1109/WIFS49906.2020.9360910.
[45] S. Suratkar, et al., Employing transfer-learning based CNN architectures to enhance the generalizability of deepfake detection, in: 2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020, 2020, https://doi.org/10.1109/ICCCNT49239.2020.9225400.
[46] S. Tariq, S. Lee, S. Woo, One detector to rule them all: towards a general deepfake attack detection framework, in: The Web Conference 2021 - Proceedings of the World Wide Web Conference, WWW, 2021, pp. 3625–3637, https://doi.org/10.1145/3442381.3449809.
[47] Eleanor Tursman, et al., Towards untrusted social video verification to combat deepfakes via face geometry consistency, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, pp. 2784–2793, https://doi.org/10.1109/CVPRW50498.2020.00335.
[48] Sheng-Yu Wang, et al., CNN-generated images are surprisingly easy to spot... for now, in: Proceedings of the IEEE/CVF Conference on Com- Puter Vision and Pattern Recognition, 2020, pp. 8695–8704.
[49] J. Wu, et al., A forensic method for DeepFake image based on face recognition, in: ACM International Conference Proceeding Series, 2020, pp. 104–108, https://doi.org/10.1145/3409501.3409544.
[50] D. Xie, et al., DeepFake detection on publicly available datasets using modified AlexNet, in: 2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020, 2020, pp. 1866–1871, https://doi.org/10.1109/SSCI47803.2020.9308428.
[51] Z. Xu, et al., Detecting facial manipulated videos based on set convolu- tional neural networks, in: Journal of Visual Communication and Image Representation 77, 2021, https://doi.org/10.1016/j.jvcir.2021.103119.
[52] J. Yang, et al., Detecting fake images by identifying potential texture difference, in: Future Generation Computer Systems 125, 2021, pp. 127–135, https://doi.org/10.1016/j.future.2021.06.043.
[53] I.-J. Yu, et al., Manipulation classification for JPEG images using multi- domain features, in: IEEE Access 8, 2020, pp. 210837–210854, https://doi.org/10.1109/ACCESS.2020.3037735.
[54] Z. Zhang, et al., Detecting manipulated facial videos: a time series solu- tion, in: Proceedings - International Conference on Pattern Recognition, 2020, pp. 2817–2823, https://doi.org/10.1109/ICPR48806.2021.9412610.
[55] Z. Zhao, P. Wang, W. Lu, Multi-layer fusion neural network for deep- fake detection, in: International Journal of Digital Crime and Forensics 13, 2021, pp. 26–39, https://doi.org/10.4018/IJDCF.20210701.oa3, 4.
[56] X. Zhou, Y. Wang, P. Wu, Detecting deepfake videos via frame serialization learning, IICSPI (2020) 391–395, https://doi.org/10.1109/IICSPI51290.2020.9332419, 2020, Proceedings of 2020 IEEE 3rd International Conference of Safe Production and Informatization.
[57] Y. Akbari, N. Almaadeed, S. Al-Maadeed, F. Khelifi, A. Bouridane, PRNU-net: a Deep Learning Approach for Source Camera Model Identification Based on Videos Taken with Smartphone, 2022.
[58] Y. Akbari, S. Al-maadeed, O. Elharrouss, F. Khelifi, A. Lawgaly, A. Bouridane, Digital forensic analysis for source video identification: a survey, Forensic Sci. Int.: Digit. Invest. 41 (2022), 301390.
[59] E. Altinisik, K. Tasdemir, H.T. Sencar, Extracting PRNU noise from H. 264 coded videos, in: 2018 26th European Signal Processing Conference (EUSIPCO), 2018, pp. 1367–1371.
[60] A. Badiye, P. Kathane, K. Krishan, Forensic Gait Analysis, 2020.
[63] I. Birch, M. Birch, N. Asgeirsdottir, The identification of individuals by observational gait analysis using closed circuit television footage: comparing the ability and confidence of experienced and non-experienced analysts, Sci. Justice 60 (1) (2020) 79–85.
[64] I. Birch, M. Birch, J. Lall, The accuracy and validity of the sheffield features of gait tool, Sci. Justice 61 (1) (2021) 72–78.
[65] I. Birch, M. Birch, L. Rutler, S. Brown, L.R. Burgos, B. Otten, M. Wiedemeijer, The repeatability and reproducibility of the sheffield features of gait tool, Sci. Justice 59 (5) (2019) 544–551.
[66] I. Birch, C. Gwinnett, J. Walker, Aiding the interpretation of forensic gait analysis: development of a features of gait database, Sci. Justice 58 (1) (2018) 78–82.
[67] I. Birch, M. Nirenberg, W. Vernon, M. Birch, Forensic Gait Analysis: Principles and Practice, CRC Press, 2020.
[68] I. Bouchrika, A survey of using biometrics for smart visual surveillance: gait recognition, in: Surveillance in Action, Springer, 2018, pp. 3–23.
[69] D. Bykhovsky, Recording device identification by ENF harmonics power analysis, Forensic Sci. Int. 307 (2020), 110100.
[71] J. Choi, C.-W. Wong, ENF signal extraction for rolling-shutter videos using periodic zero-padding, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP), 2019, pp. 2667–2671.
[72] E. Cunliffe, The legal context of forensic gait analysis: Part 1: the legal context in North America, in: Forensic Gait Analysis, CRC Press, 2020, pp. 59–70.
[73] L. de Roos, Z. Geradts, Factors that influence PRNU-based camera-identification via videos, J. Imag. 7 (1) (2021) 8.
[74] R. Deka, C. Galdi, J.-L. Dugelay, Hybrid G-PRNU: optimal parameter selection for scale-invariant asymmetric source smartphone identification, Electron. Imag. 2019 (5) (2019) 541–546.
[77] P. Ferrara, M. Iuliani, A. Piva, PRNU-based video source attribution: which frames are you using? J. Imag. 8 (3) (2022) 57.
[78] P. Ferrara, I. Sanchez, G. Draper-Gil, H. Junklewitz, L. Beslay, A MUSIC spectrum combining approach for ENF-based video timestamping, in: 2021 IEEE International Workshop on Biometrics and Forensics, (IWBF), 2021, pp. 1–6.
[79] E. Flor, R. Aygun, S. Mercan, K. Akkaya, PRNU-Based source camera identification for multimedia forensics, in: 2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science, IRI, 2021, pp. 168–175.
[80] I. Freckelton, Guarding the gait: evaluating forensic gait analysis evidence, in: Forensic Analysis-Scientific and Medical Techniques and Evidence under the Microscope, 2021.
[81] G. Frijters, Z.J.M.H. Geradts, Use of electric network frequency presence in video material for time estimation, J. Forensic Sci. 67 (3) (2022) 1021–1032.
[82] R. Garg, A. Hajj-Ahmad, M. Wu, Feasibility Study on Intra-grid Location Estimation Using Power ENF Signals, 2021. *ArXiv Preprint ArXiv:2105.00668*.
[83] H. Guan, M. Kozak, E. Robertson, Y. Lee, A.N. Yates, A. Delgado, D. Zhou, T. Kheyrkhah, J. Smith, J. Fiscus, MFC datasets: large-scale benchmark datasets for media forensic challenge evaluation, in: 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW), 2019, pp. 63–72.
[84] H. Han, Y. Jeon, B. Song, J.W. Yoon, A phase-based approach for ENF signal extraction from rolling shutter videos, IEEE Signal Process. Lett. (2022).
[85] B.C. Hosler, X. Zhao, O. Mayer, C. Chen, J.A. Shackleford, M.C. Stamm, The video authentication and camera identification database: a new database for video forensics, IEEE Access 7 (2019) 76937–76948.
[87] G. Hua, Error analysis of forensic ENF matching, in: 2018 IEEE International Workshop on Information Forensics and Security (WIFS), 2018, pp. 1–7.
[88] Z. Huang, D. Xue, X. Shen, X. Tian, H. Li, J. Huang, X.-S. Hua, 3d local convolutional neural networks for gait recognition, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 14920–14929.
[89] D. Imoto, M. Hirabayashi, M. Honma, K. Kurosawa, Enhancing the robustness of forensic gait analysis against near-distance viewing direction differences, Multimed. Tool. Appl. (2022) 1–23.
[91] G. Jackson, I. Birch, Probative value of gait analysis, in: Forensic Gait Analysis, CRC Press, 2020, pp. 171–188.
[92] P.R.M. Júnior, L. Bondi, P. Bestagini, A. Rocha, S. Tubaro, A prnu-based method to expose video device compositions in open-set setups, in: IEEE International Conference on Image Processing (ICIP), 2019, pp. 96–100, 2019.
[93] H.D. Kelly, Forensic Gait Analysis, CRC Press, 2020.
[94] N.-T. Kim, S.-J. Moon, M.-S. Park, K.-M. Lee, K.-J. Choi, W.-Y. Choi, K.-H. Sung, Prevalence of gait features in healthy adolescents and adults, Kor. J. Legal Med. 45 (1) (2021) 27–33.
[95] E.K. Kouokam, A.E. Dirik, PRNU-based source device attribution for YouTube videos, Digit. Invest. 29 (2019) 91–100.

[96] J. Kwon, Y. Lee, J. Lee, Comparative study of markerless vision-based gait analyses for person Re-identification, Sensors 21 (24) (2021) 8208.
[97] A. Lawgaly, F. Khelifi, A. Bouridane, S. Al-Maaddeed, Sensor pattern noise estimation using non-textured video frames for efficient source smartphone identification and verification, in: 2021 International Conference on Computing, Electronics & Communications Engineering, (ICCECE), 2021, pp. 19–24.
[98] A. Lawgaly, F. Khelifi, A. Bouridane, S. Al-Maaddeed, Y. Akbari, PRNU estimation based on weighted averaging for source smartphone video identification, in: 2022 8th International Conference on Control, Decision and Information Technologies 1, (CoDIT), 2022, pp. 75–80.
[99] A. Lawgaly, F. Khelifi, A. Bouridane, S. Al-maadeed, Y. Akbari, Three dimensional denoising filter for effective source smartphone video identification and verification, in: 2022 7th International Conference on Machine Learning Technologies, (ICMLT), 2022, pp. 124–130.
[100] J. Li, B. Ma, C. Wang, Extraction of PRNU noise from partly decoded video, J. Vis. Commun. Image Represent. 57 (2018) 183–191.
[101] R.R. López, E.A. Luengo, A.L.S. Orozco, L.J.G. Villalba, Digital video source identification based on container's structure analysis, IEEE Access 8 (2020) 36363–36375.
[103] F. Lugstein, S. Baier, G. Bachinger, A. Uhl, PRNU-based deepfake detection, in: Proceedings of the 2021 ACM Workshop on Information Hiding and Multimedia Security, 2021, pp. 7–12.
[104] I. Macoveciuc, C.J. Rando, H. Borrion, Forensic gait analysis and recognition: standards of evidence admissibility, J. Forensic Sci. 64 (5) (2019) 1294–1303.
[105] I. Macoveciuc, C.J. Randob, R. Morgana, E. Cremac, The Utility of Gait in Forensic Human Identification: an Empirical Investigation Using Biomechanical and Anthropological Principals, Defence and Security Doctoral Symposium, 2020, 2020.
[106] Z. Mahfouf, H.F. Merouani, I. Bouchrika, N. Harrati, Investigating the use of motion-based features from optical flow for gait recognition, Neurocomputing 283 (2018) 140–149.
[107] L. Maiano, I. Amerini, L. Ricciardi Celsi, A. Anagnostopoulos, Identification of social-media platform of videos through the use of shared features, J. Imag. 7 (8) (2021) 140.
[108] A. Majeed, A.K. Chong, Study of CCTV Footage Based on Lower-Limb Gait Measure for Forensic Application, in: 2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC), 2021, pp. 160–164.
[110] S. Mandelli, F. Argenti, P. Bestagini, M. Iuliani, A. Piva, S. Tubaro, A modified Fourier-Mellin approach for source device identification on stabilized videos, in: 2020 IEEE International Conference on Image Processing, ICIP), 2020, pp. 1266–1270.
[111] S. Mandelli, P. Bestagini, S. Tubaro, D. Cozzolino, L. Verdoliva, Blind detection and localization of video temporal splicing exploiting sensor-based footprints, in: 2018 26th European Signal Processing Conference, EUSIPCO), 2018, pp. 1362–1366.
[113] O. Mayer, B. Hosler, M.C. Stamm, Open set video camera model verification, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP), 2020, pp. 2962–2966.
[114] S. Mercan, M. Cebe, R.S. Aygun, K. Akkaya, E. Toussaint, D. Danko, Blockchain-based video forensics and integrity verification framework for wireless Internet-of-Things devices, Security and Privacy 4 (2) (2021) e143.
[115] P.P. Min, S. Sayeed, T.S. Ong, Gait recognition using deep convolutional features, in: 2019 7th International Conference on Information and Communication Technology, (ICoICT), 2019, pp. 1–5.
[116] Z. Mu, F.M. Castro, M.J. Mar\'\in-Jiménez, N. Guil, Y.-R. Li, S. Yu, Resgait: the real-scene gait dataset, in: 2021 IEEE International Joint Conference on Biometrics, (IJCB), 2021, pp. 1–8.
[117] D. Nagothu, Y. Chen, A. Aved, E. Blasch, Authenticating video feeds using electric network frequency estimation at the edge, EAI Endorsed Trans. Secur. Saf. 7 (24) (2021) e4, e4.
[118] D. Nagothu, Y. Chen, E. Blasch, A. Aved, S. Zhu, Detecting Malicious False Frame Injection Attacks on Video Surveillance at the Edge Using Electrical Network Frequency Signals, 2019.
[119] D. Nagothu, J. Schwell, Y. Chen, E. Blasch, S. Zhu, A study on smart online frame forging attacks against video surveillance system, Sens. Syst. Space Appl. XII (2019) 176–188, 11017.
[120] K.K. Nagwanshi, Cyber Forensic Review of Human Footprint and Gait-Based System for Personal Identification in Crime Scene Investigation, 2018.
[121] M. Nielsen, N. Lynnerup, P.K. Larsen, Forensic anthropological video-based cases at the department of forensic medicine, university of Copenhagen: a 10-year retrospective review, Scandinavian J. Forensic Sci. 25 (1) (2019) 9–13.
[122] M. Nirenberg, The history of the use of gait analysis as evidence, in: Forensic Gait Analysis, CRC Press, 2020, pp. 19–38.
[123] M. Nirenberg, W. Vernon, I. Birch, A review of the historical use and criticisms of gait analysis evidence, Sci. Justice 58 (4) (2018) 292–298.
[125] H. Qin, Z. Chen, Q. Guo, Q.M.J. Wu, M. Lu, Rpnet: gait recognition with relationships between each body-parts, IEEE Trans. Circ. Syst. Video Technol. 32 (5) (2021) 2990–3000.
[126] H. Razak, A.A. Almisreb, N.M. Tahir, Detection of anomalous gait as forensic gait in residential units using pre-trained convolution neural networks, Future Inf. Commun. Conf. (2020) 775–793.
[127] S. Reel, I. Birch, J. Saxelby, S. Reidy, Code of Practice for Forensic Gait Analysis, 2019.
[129] X. Ren, S. Hou, C. Cao, X. Liu, Y. Huang, Progressive Feature Learning for Realistic Cloth-Changing Gait Recognition, 2022. *ArXiv Preprint ArXiv: 2207.11720*.
[130] A. Sakata, Y. Makihara, N. Takemura, D. Muramatsu, Y. Yagi, How confident are you in your estimate of a human age? uncertainty-aware gait-based age estimation by label distribution learning, in: 2020 IEEE International Joint Conference on Biometrics, (IJCB), 2020, pp. 1–10.
[131] A. Sakata, N. Takemura, Y. Yagi, Gait-based age estimation using multi-stage convolutional neural network, IPSJ Trans. Comput. Vision Appl. 11 (1) (2019) 1–10.
[132] L. Satchell, Psychology of perceptual error in forensic practice, in: Forensic Gait Analysis, CRC Press, 2020, pp. 159–170.
[133] D. Seckiner, The Development and Testing of a Forensic Interpretation Framework for Use on Anthropometric and Morphological Data Collected during Stance and Gait, 2021.
[134] D. Seckiner, X. Mallett, C. Roux, D. Meuwly, P. Maynard, Forensic image analysis–CCTV distortion and artefacts, Forensic Sci. Int. 285 (2018) 77–85.
[135] H. Sharma, N. Kanwal, R.S. Batth, An ontology of digital video forensics: classification, research gaps & datasets, in: 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE), 2019, pp. 485–491.
[136] B. Singh, K. Krishan, K. Kaur, T. Kanchan, Estimation of body weight from the base of gait and the area swept in one stride—forensic implications, Egypt. J. Food Sci. 8 (1) (2018) 1–7.
[138] M. Solon, Face coverings do not mask your gait: wearing face masks or face coverings are now mandatory in numerous environments, but miscreants and protestors hoping to take advantage of the anonymity masks give may find themselves down on their luck as surveillance has moved up a stride, as Mark Solon explains, Podiatry Review 77 (4) (2020) 28–31.
[139] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, Y. Yagi, Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition, IPSJ Trans. Comput. Vision Appl. 10 (1) (2018) 1–14.
[140] M. Uddin, T.T. Ngo, Y. Makihara, N. Takemura, X. Li, D. Muramatsu, Y. Yagi, others, The ou-isir large population gait database with real-life carried object and its performance evaluation, IPSJ Trans. Comput. Vision Appl. 10 (1) (2018) 1–11.
[141] N.M. van Mastrigt, K. Celie, A.L. Mieremet, A.C.C. Ruifrok, Z. Geradts, Critical review of the use and scientific basis of forensic gait analysis, Forensic Sci. Res. 3 (3) (2018) 183–193.
[142] L. Verdoliva, Media forensics and deepfakes: an overview, IEEE J. Selected Topics Signal Process. 14 (5) (2020) 910–932.
[143] M.M. Wiedemeijer, Evaluation of Expert Observations of Gait in Forensics, 2019.
[144] H. Wu, J. Tian, Y. Fu, B. Li, X. Li, Condition-aware comparison scheme for gait recognition, IEEE Trans. Image Process. 30 (2020) 2734–2744.
[145] J. Xiao, S. Li, Q. Xu, Video-based evidence analysis and extraction in digital forensic investigation, IEEE Access 7 (2019) 55432–55442.
[146] R. Xu, D. Nagothu, Y. Chen, Decentralized video input authentication as an edge service for smart cities, IEEE Consumer Electronics Magazine 10 (6) (2021) 76–82.
[147] W.-C. Yang, J. Jiang, C.-H. Chen, A fast source camera identification and verification method based on PRNU analysis for use in video forensic investigations, Multimed. Tool. Appl. 80 (5) (2021) 6617–6638.
[148] P. Yu, Z. Xia, J. Fei, Y. Lu, A survey on deepfake video detection, IET Biom. 10 (6) (2021) 607–624.
[149] L. Zhang, T. Qiao, M. Xu, N. Zheng, S. Xie, Unsupervised learning-based framework for deepfake video detection, IEEE Trans. Multimed. (2022).