International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 12 August 2023 **Revised:** 30 September 2023 **Accepted:** 15 October 2023*

Deception/Truthful Prediction Based on Facial Feature and Machine Learning Analysis

Twisha Patel[1], Dr. Daxa Vekariya[2]
[1]Research Scholor
Engineering and Technology,
Parul University, Vadodara,
Gujarat, India
twisapatel@gmail.com
[2]Faculty
Engineering and Technology,
Parul University, Vadodara,
Gujarat, India
daxa.vekariya18436@paruluniversity.ac.in

Abstract— The Automatic Deception detection refers to the investigative practices used to determine whether person is telling you Truth or lie. Automatic deception detection has been studied extensively as it can be useful in many real-life scenarios in health, justice, and security systems. Many psychological studies have been reported for deception detection. Polygraph testing is a current trending technique to detect deception, but it requires human intervention and training. In recent times, many machine learning based approaches have been applied to detect deceptions. Various modalities like Thermal Imaging, Brain Activity Mapping, Acoustic analysis, eye tracking. Facial Micro expression processing and linguistic analyses are used to detect deception. Machine learning techniques based on facial feature analysis look like a promising path for automatic deception detection. It also works without human intervention. So, it may give better results because it does not affect race or ethnicity. Moreover, one can do covert operation to find deceit using facial video recording. Covert Operation may capture the real personality of deceptive persons. By making combination of various facial features like Facial Emotion, Facial Micro Expressions and Eye blink rate, pupil size, Facial Action Units we can get better accuracy in Deception Detection.

Keywords- Facial Feature; Dlib; Support vector machine (SVM); K-nearest Neighbours; Random Forest; Decision Tree; Extra Tree; and Naive Baye.

I. INTRODUCTION

Deception is the action of deceiving someone or hiding something. Now a day’s deception is widespread in mankind. It has many from but is mainly seen as high-stake environments and low stake deception. Developing a method which can detect deception behavior has become important. As detecting deception has become more challenging [1]. When a speaker makes a remark and the imposter interprets that speech as a major conclusion, high stack deception occurs. For example, in judicial cases, a false declaration might result in a guilty person being released without charges. Low stake deception has no or minimal significance for the impostor. e.g., social media posted, online reviews [2].

Deception detection is used in many areas in real life such as health, justice, and security systems. But as now a day’s deception detection is mainly used in criminal investigations, where criminal will try to make a false statement as he/she do not want to face the punishment. The ability of detecting deception by humans is 54% without using any specialized machine [1]. Therefore, building a system which as deception detection has become important. There are many psychological techniques which detected the deception which are Polygraph testing, Questioning, and testing techniques, Voice stress, Brain scan etc. [2]. Polygraph testing is more popular for detecting deception [2]. But this all-psychological technique required specialized machines and required human expertise to detect deception. Therefore, use of it in the real world is less. Researchers have proposed many machines learning for detecting deception such Thermal Imaging, Brain Activity Mapping, Acoustic analysis, eye tracking. They also develop many methods to use facial expressions. Facial micro expression is used in detecting deception, but it is difficult to detect as it stays for short duration, which is 500ms, other has given as 250ms, 330ms and 200ms [3] and it is hard to detect by naked eyes. However, facial features play an important for detecting deception in nonverbal seen and does not require any person.

In this paper, suggested the deception detection using facial video recording. We have combined various facial features such Facial Emotion, Facial Micro Expressions and Eye blink rate, pupil size, Facial Action Units for detecting deception. The remaining research paper consists of. Section 2 contents the information of the related Literature, Furthermore Section 3 contents suggested methods, and Moreover Section 4 shows outcome and analysis. Finally, Section 5 has the conclusions of this research work and projection.

II. RELETED WORKS

In [1] Leena Mathur et al. has introduced the Automated Multimodal Deception Detection using features from audio, video, vocal. Taking advantage of the most recent developments in emotion identification in the wild, they used a state-of-the-art deep neural network trained on the A database to extract continuous models of speakers' faces' valence and arousal levels. There, they put their model through its paces using both unimodal and multimodal SVM-based.

The hybrid network employed by C.Lic et al. for feature extraction in facial emotion recognition is described in [3]. (FER). The proposed network is a combination of a Spatial Attention Convolutional Neural Network and a series of Long Short-term Memory networks that make use of the attention mechanism that each possesses. Three datasets (FER2013, CKC, and JAFFE) were employed. But the drawback of this was that the proposed network was complex and difficult to implement in real-world applications.

In this [4] the Gullapalli, A. R. et al. has tested the idea of automated detection of eye blink and blink frequency dynamics in a forensic sample might predict trait degrees of dishonesty and devised a unique automated image analysis method based on hidden Markov machines, in which an underlying hidden process is assumed to impact the visible, measurable process [4]. But in this, the experiment was done only on adult male samples and detection of input video is whether deceptive or not was only done by one feature i.e blinking of eyes [4].

In [5] Karnati, M. et al. has used a hybrid strategy, combining interaction and non-contact procedures. They have used 2D images, EEG signals and LieNet model. Dataset were MU3D and RL. In [6] M. Monaro et al. have used latency-based techniques. The researchers wanted to see if merging a choice reaction time (RT) methodology with a unique open questionnaire might help them detect identity liars. The Inverse Efficiency Score to Unexpected Questions is the highly informative trait in differentiating deceptive from truth-speakers, according to the findings. In this bandwidth was an issue [6].

In [7] X. Shu et al. has given the description and recognition of facial texture, the ED-LBP stand for equilibrium difference local binary pattern is a unique texture descriptor. They have used LBP, ED-LBP histograms and SVM classifier. The dataset which they used were Replay-Attack, Replay-Mobile, OULU-NPU and CASIA FASD [7]. In [8] C. Dalvi et al. it was a survey on the Facial Recognition, and they have offered a thorough assessment of AI-based FER and mentioned different methodology, all datasets, feature extraction techniques, algorithms, and current achievements in facial expression identification [8].

In [9] this the Zhang et al. has described fusing multiple weak classifiers to detect common facial expressions. Using an updated weighted mean value technique, it combines each classifier's prediction results and suggests an expression feature extraction method based on key point detection and dataset FER 2013 and CK+. In [10] Feng Ding et al. has suggested a GAN model that acts as a forensics deterrent. It has a unique architecture that includes additional overseeing modules to improve image visual quality. They have used a deepfake method called as black-box anti-forensics.

In [11] N. Samadiani et al. has discussed three significant challenges in unrestricted real-world circumstances, including lighting variance, head posture, and subject-dependence, that may not be solved through FER image/video analysis. There concentrate on sensors that could supply additional information to FER systems, allowing them to identify emotion from both pictures that are static and video sequences. In [12] D. K. Jain et al. has descried the methods to sort each image into one of six different face expression categories. Convolution layers and deep residual blocks by single deep convolutional Neural Networks were used to build a model and two datasets CK+ and JAFFE were used.

In [13] D. Avola et al. has detected deceit in RGB videos using they method. From video frames the authors have detected AU stands for Action Units and by using SVM they have detected true speaker or deceptive speaker from the given input of participants. In [14] V. Gupta et al has given a novel multimodal dataset is offered that contains information for deception detection using different modalities, including data from video, audio, EEG, and eye. The dataset given here was gathered in a realistic environment and includes 35 distinct respondents who contributed 325 data with labelled, in which 163 were truth and 162 were falsehood.

In [15] T. Baltrusaitis et al. has developed the toolkit that can run in real-time and can be powered by a basic webcam without the need for any other hardware. And for research purposes, the OpenFace 2.0 source code for training and running models is publicly accessible. In [16] V. Pérez-Rosas has created a real-world trial dataset which contains clips from public court cases. For testing the given input as truthful and deceptive the researcher has made a multimodal deception detection system. Researchers combined verbal and nonverbal modalities.

III. PROPOSED METHODOLOGY

A. Video

Input is given as video. To check whether a given input is Deceptive or Not. Once the input the given using open CV the resizing is done and from the given input video it is converted into frames. We have tried on 400 frames for accuracy results.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 12 August 2023 **Revised:** 30 September 2023 **Accepted:** 15 October 2023*

B. Face Detection

After the frame are it will go face detection and which done by using HOG and SVM Linear. In which HOG is used to extract the feature from the image and based on this data SVM Linear will detect the face. After detecting the face, the Landmark are generated using DLIB.

- HOG: The abbreviation for Histogram of Oriented Gradients is HOG. In computer vision and image processing, feature extraction is used for object recognition. To reliably detect objects, HOG relies on the local intensity gradient distribution and the direction of edges to define the object form [17].
- Support Vector Machine Algorithm (SVM): The goal is to determine the best line for future segmentation of n-dimensional space into classes, making it possible to categories new data points more accurately. The best line, also called a hyperplane [18].
- Dlib Facial Landmarks: It's a pre-trained detector for facial landmarks, and it gives 68 coordinates (x, y) of the facial points on a person's face [19]. It detects Eyes, Eyebrows, Nose, Mouth, Jawline.

C. Facial Features

Furthermore, Ones the 68 landmarks are detected from that facial feature are taken such as Facial Emotions, Facial Unit Action, Facial Micro Expressions and Eye Movement.

- Facial Emotions: Facial Emotion is detected by SVM and CK+ dataset.
- CK+ Dataset: The database is managed in a controlled laboratory environment and contains an evaluation of the FER system. The 593 videos available on CKC cover a wide range of topics. The time span might be anything between 10 and 60 seconds. Three hundred twenty-seven sequences from 118 people were found using FACS, and these sequences contained seven basic expressions: fear, anger, surprise, disgust, disgust, pleasure, sadness, and contempt.
- Facial Unit Action: The Face Action Coding Approach (FACS) is a comprehensive approach for describing all clearly identifiable facial movement that is based on anatomical principles. It breaks down facial emotions into Action Units, which are little chunks of muscle activity (AUs) [21].
- Facial Micro Expression: Facial micro-expressions are very brief, spontaneous facial expressions that people make when they are trying to hide an emotion, whether consciously or unconsciously. It stays for short duration, which is 500ms, other has given as 250ms, 330ms and 200ms [12].
- Eye Movement: Eye Movement are the actions taken by eyes such as eye close, eye open, repeatedly eye open/close etc.

D. Dataset:

We have used “Real-life Trial Data” dataset. A multimodal dataset of real-life deception: deceptive and honest trial testimony that were personally transcribed and annotated. There are 121 videos in all, with 61 misleading and 60 true trial pieces. The videos in the sample are 28.0 seconds long on average. The deceptive clips have an average video time of 27.7 seconds, whereas the genuine clips have an average video duration of 28.3 seconds. There are 21 distinct female speakers and 35 unique male speakers in the data, with ages ranging from 16 to 60 [16].

E. Machine Learner:

- Naive Bayes: It is following Bayes' Theorem. A method for figuring out conditional probability based on previous information and the naïve assumption that each feature is independent of the others. The most significant benefit of Naive Bayes is that, whereas other machine learning algorithms require a substantial quantity of training data, it performs admirably even when the training data is limited [18].
- K-nearest Neighbors: The unknown data is labelled by the nearest observed data point, and which is calculated by the distance between 2 points. The k nearest neighbor approach is used to represent each data point in an n-dimensional space defined by n attributes [18].
- Decision Tree: In a hierarchical method, a decision tree creates tree branches, each of which can be thought of as an if-else expression. The branches are created by breaking down the dataset into subgroups based on the most important attributes. The leaves of the decision tree are where the ultimate categorization occurs [18].
- Random Forest: It is a set of multiple decision trees arranged in a random pattern. It's a form of ensemble approach that combines the findings of several predictors. It also adopts a bagging approach, in which a random sample of the original dataset is used to train each tree. The trees are voted on by the majority. [18].
- Support vector machine (SVM): It determines the optimal method to classify the data based on its location in reference to a positive/negative class boundary. The hyperplane is a boundary that maximizes the distance between data points of different classifications [18].
- Extra Trees: This class implements a Meta estimator that employs averaging to increase predicted accuracy and control over-fitting by fitting several randomized decision trees on various sub-samples of the dataset [22].

F. Parameters

TABLE I. CONFUSION MATRIX EXAMPLE

Negative	Positive

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *12 August 2023* ***Revised:*** *30 September 2023* ***Accepted:*** *15 October 2023*

Negative	4	2
Positive	1	5

- Accuracy: It is ratio of sum of positively classified instances to total of number instances of data [23].

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN} \quad (1)$$

In this case, the accuracy is calculated as follows: Accuracy = (4 + 5) / (4 + 2 + 1 + 5) = (0.75).

- Precision: To finding total expected positive observations is the ratio of the completely expected positive [23].

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Given this, we can calculate that the accuracy of the given example is 0.80, or 4 / (4 + 1) / (4 + 1).

- Recall: All observations in the class is compared with the correctly predicted positive observations in actual class and the proportion this is known as recall.[23].

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Accordingly, the calculated value of recall in this scenario is (four) / (four plus two) = (0.67).

- F1-Score: Both false positives and false negatives are considered in this score and which the weighted average of Precision and Recall [23].

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

So, the F1-score in this case is calculated as 2*(0.80 * 0.67)/(0.80 + 0.67) = 0.7292.

IV. ANALYSIS OF RESULTS

```
<bound method NDFrame.head of                   id  OtherGestures  Smile  ...  downHands  upHands      class
0      trial_lie_001.mp4           1      0  ...          0        0  deceptive
1      trial_lie_002.mp4           1      0  ...          0        0  deceptive
2      trial_lie_003.mp4           1      0  ...          0        0  deceptive
3      trial_lie_004.mp4           1      0  ...          0        0  deceptive
4      trial_lie_005.mp4           1      0  ...          0        0  deceptive
..                   ...         ...    ...  ...        ...      ...        ...
116  trial_truth_056.mp4           1      0  ...          0        1   truthful
117  trial_truth_057.mp4           1      0  ...          0        0   truthful
118  trial_truth_058.mp4           1      0  ...          0        0   truthful
119  trial_truth_059.mp4           0      0  ...          0        0   truthful
120  trial_truth_060.mp4           0      1  ...          0        0   truthful

[121 rows x 41 columns]>
```

Figure 2. Dataset Reading

```
<bound method NDFrame.head of      Smile  Laugh  Scowl  ...  lipsUp  lipsRetracted  lipsProtruded
0        0      0      0  ...       0              0              0
1        0      0      0  ...       0              0              0
2        0      0      0  ...       0              0              0
3        0      0      0  ...       0              0              0
4        0      0      0  ...       0              0              0
..     ...    ...    ...  ...     ...            ...            ...
116      0      0      0  ...       0              0              0
117      0      0      0  ...       0              0              0
118      0      0      0  ...       1              0              0
119      0      0      1  ...       0              0              0
120      1      0      0  ...       1              0              0

[121 rows x 20 columns]>
<bound method NDFrame.head of 0      deceptive
1      deceptive
2      deceptive
3      deceptive
4      deceptive
         ...
116     truthful
117     truthful
118     truthful
119     truthful
120     truthful
Name: class, Length: 121, dtype: object>
K-Nearest Neighbors
```

Figure 3. Feature Extracted

We have selected 20 features for the original Dataset. 20 feature includes Smile, Laugh, Scowl, otherEyebrowMovement, Frown, Raise, OtherEyeMovements, Close-R, X-Open, Close BE, gazeInterlocutor, gazeDown, gazeUp, otherGaze, openMouth, closeMouth, lipsDown, lipsUp, lipsRetracted, lipsProtruded.eature selection.

```
Name: class, Length: 121, dtype: object>
K-Nearest Neighbors
              precision    recall  f1-score   support

   deceptive       0.83      0.83      0.83         6
    truthful       0.86      0.86      0.86         7

    accuracy                           0.85        13
   macro avg       0.85      0.85      0.85        13
weighted avg       0.85      0.85      0.85        13

[[5 1]
 [1 6]]
-------------------------------------------------------
```

Figure 4. KNN

```
-------------------------------------------------------
Liner SVM
              precision    recall  f1-score   support

   deceptive       0.86      1.00      0.92         6
    truthful       1.00      0.86      0.92         7

    accuracy                           0.92        13
   macro avg       0.93      0.93      0.92        13
weighted avg       0.93      0.92      0.92        13

[[6 0]
 [1 6]]
-------------------------------------------------------
```

Figure 5. SVM

```
-------------------------------------------------------
Decision Tree
              precision    recall  f1-score   support

   deceptive       0.83      0.83      0.83         6
    truthful       0.86      0.86      0.86         7

    accuracy                           0.85        13
   macro avg       0.85      0.85      0.85        13
weighted avg       0.85      0.85      0.85        13

[[5 1]
 [1 6]]
-------------------------------------------------------
```

Figure 6. Decision Tree

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 12 August 2023 **Revised:** 30 September 2023 **Accepted:** 15 October 2023*

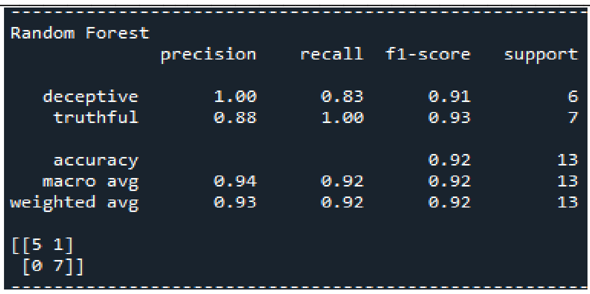

Figure 7. Random Forest

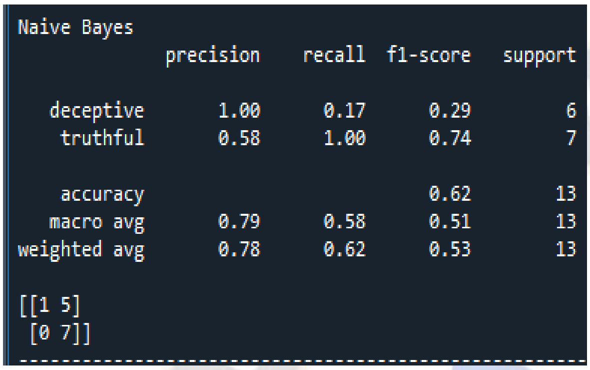

Figure 8. Naive Bayes

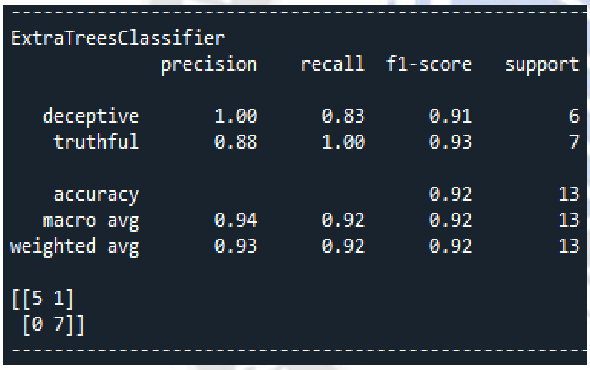

Figure 9. Extra Tree

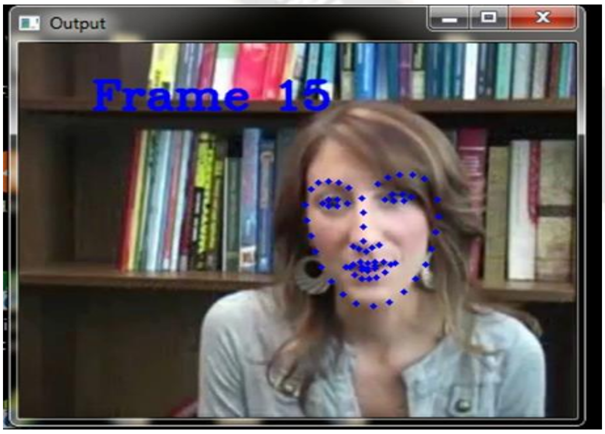

Figure 10. Video

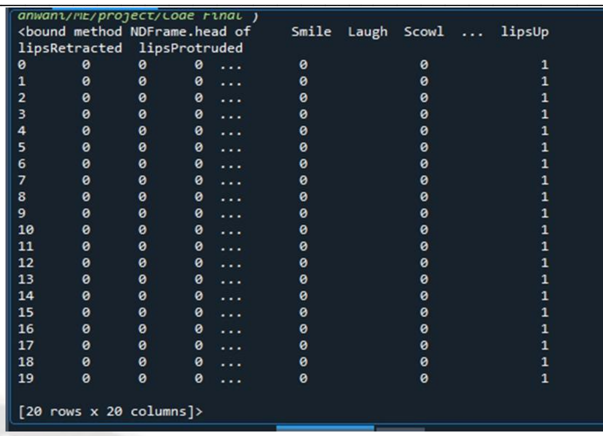

Figure 11. Truth Dataset

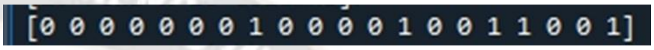

Figure 12. Average of all frames

TABLE II. PERFORMANCE ANALYSIS

Models	Precision	Recall	F1-Score	Accuracy
KNN	0.85	0.85	0.85	0.85
SVM	0.93	0.93	0.92	0.92
DT	0.85	0.85	0.85	0.85
RF	0.94	0.92	0.92	0.92
NB	0.79	0.58	0.51	0.62
ET	0.94	0.92	0.92	0.92

Figure 13. Performance graph

V. CONCLUSION

Deception detection has gained widespread importance in various real-life contexts due to its utility in discerning truth from falsehood. Human perception alone can often struggle to detect deception accurately. Consequently, we have devised a model to effectively identify deception in given inputs. Leveraging Dlib, we have successfully pinpointed 68 facial landmarks and employed SVM for emotion recognition. During testing, our model excelled, with the best results achieved using classifiers like RF, ET, and SVM. Looking forward, our future efforts will focus on incorporating additional body features into our model to enhance its accuracy and overall performance in detecting deception.

REFERENCES

[1] Jr. Charles F. Bond and Bella M. DePaulo. "Accuracy of deception judgments. Personality and Social Psychology Review," 10(3):214–234, 2006.

[2] Leena Mathur, Maja J Matarić, "Introducing Representations of Facial Affect in Automated Multimodal Deception Detection," Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI) , DOI: 10.1145/3382507.3418864.

[3] C. Liu, K. Hirota, J. Ma, Z. Jia and Y. Dai, "Facial Expression Recognition Using Hybrid Features of Pixel and Geometry," in IEEE Access, vol. 9, pp. 18876-18889, 2021, doi: 10.1109/ACCESS.2021.3054332.

[4] Gullapalli, A. R., Anderson, N. E., Yerramsetty, R., Harenski, C. L., & Kiehl, K. A. (2021),"In the blink of an eye: Quantitative blink dynamics predict deceptive personality traits in forensic interviews", in Personality and Individual Differences, 176, 110764. doi:10.1016/j.paid.2021.110764.

[5] Karnati, M., Seal, A., Yazidi, A., & Krejcar, O. (2021),"LieNet: A Deep Convolution Neural Networks Framework for Detecting Deception", in IEEE Transactions on Cognitive and Developmental Systems, 1 1. doi:10.1109/tcds.2021.3086011.

[6] M. Monaro, I. Zampieri, G. Sartori, P. Pietrini, and G. Orrù, "The detection of faked identity using unexpected questions and choice reaction times," Psychol. Res., vol. 85, no. 6, pp. 2474–2482, 2021.

[7] X. Shu, H. Tang, and S. Huang, "Face spoofing detection based on chromatic ED-LBP texture feature," Multimed. Syst., vol. 27, no. 2, pp. 161–176, 2021.

[8] C. Dalvi, M. Rathod, S. Patil, S. Gite and K. Kotecha, "A Survey of AI-Based Facial Emotion Recognition: Features, ML & DL Techniques, Age-Wise Datasets and Future Directions," in IEEE Access, vol. 9, pp. 165806-165840, 2021, doi: 10.1109/ACCESS.2021.3131733.

[9] Zhang, Chuanjie, and Changming Zhu. "Multiple classifiers fusion for facial expression recognition." Granular Computing 7.1 (2022): 171-181.

[10] Ding, F., Zhu, G., Li, Y., Zhang, X., Atrey, P. K., & Lyu, S., "Anti-Forensics for Face Swapping Videos via Adversarial Training", in IEEE Transactions on Multimedia, 1-1. doi:10.1109/tmm.2021.3098422.

[11] N. Samadiani et al., "A review on automatic facial Expression Recognition systems assisted by multimodal sensor data," Sensors (Basel), vol. 19, no. 8, p. 1863, 2019.

[12] D. K. Jain, P. Shamsolmoali, and P. Sehdev, "Extended deep neural network for facial emotion recognition," Pattern Recognit. Lett., vol. 120, pp. 69–74, 2019.

[13] D. Avola, L. Cinque, G. L. Foresti, and D. Pannone, "Automatic Deception Detection in RGB videos using Facial Action Units," in Proceedings of the 13th International Conference on Distributed Smart Cameras, 2019.

[14] V. Gupta, M. Agarwal, M. Arora, T. Chakraborty, R. Singh, and M. Vatsa, "Bag-of-lies: A multimodal dataset for deception detection," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 83–90.

[15] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency, "OpenFace 2.0: Facial behavior analysis toolkit," in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018, pp. 59–66.

[16] V. Pérez-Rosas, M. Abouelenien, R. Mihalcea, and M. Burzo, "Deception Detection using Real-life Trial Data," in Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, 2015.

[17] "Face detection using dlib HOG", Srujan Jack. Online Available: https://medium.com/mlcrunch/face-detection using-dlib-hog-198414837945.

[18] D. Gong, "Top 6 machine learning algorithms for classification," Towards Data Science, 23-Feb-2022. [Online]. Available: https://towardsdatascience.com/top-machine-learning-algorithms-for-classification-2197870ff501.

[19] I. José, "Facial mapping (landmarks) with Dlib + python," Towards Data Science, 29-Jun-2018. [Online]. Available: https://towardsdatascience.com/facial-mapping-landmarks-with-dlib-python-160abcf7d672.

[20] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

[21] "Facial action coding system," Paul Ekman Group, 21-Jun-2013. [Online]. Available: https://www.paulekman.com/facial-action-coding-system/.

[22] P. Geurts, D. Ernst., and L. Wehenkel, "Extremely randomized trees", Machine Learning, 63(1), 3-42, 2006.

[23] Accuracy, precision, recall & F1 score: Interpretation of performance measures," Exsilio Blog, 09-Sep-2016. [Online]. Available: https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/. [Accessed: 28-Apr-2022].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 12 August 2023 **Revised:** 30 September 2023 **Accepted:** 15 October 2023*

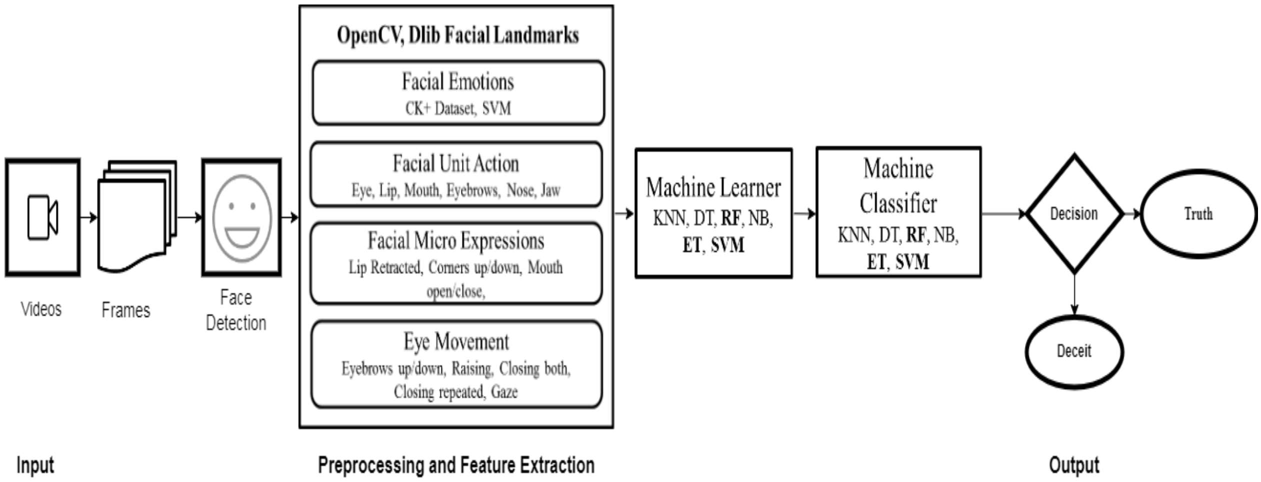

Figure 1. Proposed System