

# SOCIAL MEDIA ANALYTICS WITH DETECTION OF MODIFIED CONTENT

<sup>1</sup>Dr.Sunil Bhutada, <sup>2</sup>Mr. Punna Mahesh Kumar, <sup>3</sup>Dr.K.Kranthi Kumar,

<sup>4</sup>Sirikonda Neelima, <sup>5</sup>Zuha Rahman, <sup>6</sup>Kummari sruthilaya

<sup>1</sup>Professor & HOD, Sreenidhi Institute of Science and Technology, Dept of IT, Hyderabad.

<sup>2</sup>Assistant Professor, Sreenidhi Institute of Science and Technology, Dept of IT, Hyderabad.

<sup>3</sup>Associate Professor, Sreenidhi Institute of Science and Technology, Dept of IT, Hyderabad.

<sup>4</sup>Dept of IT, UG Student, Sreenidhi Institute of Science and Technology, Hyderabad.

<sup>5</sup>Dept of IT, UG Student, Sreenidhi Institute of Science and Technology, Hyderabad.

<sup>6</sup>Dept of IT,UG Student, Sreenidhi Institute of Science and Technology, Hyderabad.

<sup>1</sup>[sunil.b@sreenidhi.edu.in](mailto:sunil.b@sreenidhi.edu.in), <sup>2</sup>[mahesh.p@sreenidhi.edu.in](mailto:mahesh.p@sreenidhi.edu.in) , <sup>3</sup>[kranthi.k@sreenidhi.edu.in](mailto:kranthi.k@sreenidhi.edu.in)

<sup>4</sup>[22311A12A5@it.sreenidhi.edu.in](mailto:22311A12A5@it.sreenidhi.edu.in) <sup>5</sup>[22311A12C7@it.sreenidhi.edu.in](mailto:22311A12C7@it.sreenidhi.edu.in)

<sup>6</sup>[22311A12C8@it.sreenidhi.edu.in](mailto:22311A12C8@it.sreenidhi.edu.in)

## ABSTRACT

The fast growth of social media has caused a big increase in altered images and false text, making it hard to know if online information is real. Most current systems check either images or text separately, which makes it hard to catch tricky fake posts. This new system uses Error Level Analysis (ELA) to spot image changes and Natural Language Processing (NLP) to find signs of fake news or clickbait in text descriptions. The results from both image and text checks are combined to decide if the content is genuine. This mixed method works better than old ways and could get even better with models like Convolutional Neural Networks (CNN) and BERT in the future. The limitation include that the system uses rule-based NLP for text, which might not fully understand complex meanings in captions. Also, ELA alone might miss very advanced image changes or deepfake content.

**Keywords:** Image Forgery Detection, Social Media Content Analysis, Error Level Analysis(ELA), Natural language Processing(NLP), Multi-model analysis, fake content detection & authentication.

## 1. INTRODUCTION.

ISSN:2250-3676

## Motivation:

With the fast growth of social media, it's become a big problem that altered images and false captions are spreading quickly. Digital content can be changed using powerful tools, making it hard for people to tell what's real and what's fake. This has caused a lot of false information to spread, which can sway public opinion, cause confusion, and even lead to social and political problems. Most current systems look at images or text on their own. But in reality, fake content usually involves both pictures and words. An image might look real but have a wrong caption, or a caption might be true but the image is fake. Because of this, there's a need for a better way that checks both image and text together. Another key reason for this project is the need for a simple, fast, and real-time solution. Many top detection tools use complex deep learning models that need big data and powerful computers. This makes them hard to use in everyday situations. That's why there's a need for a lighter system that can work well without needing expensive equipment.

## 2. OBJECTIVES

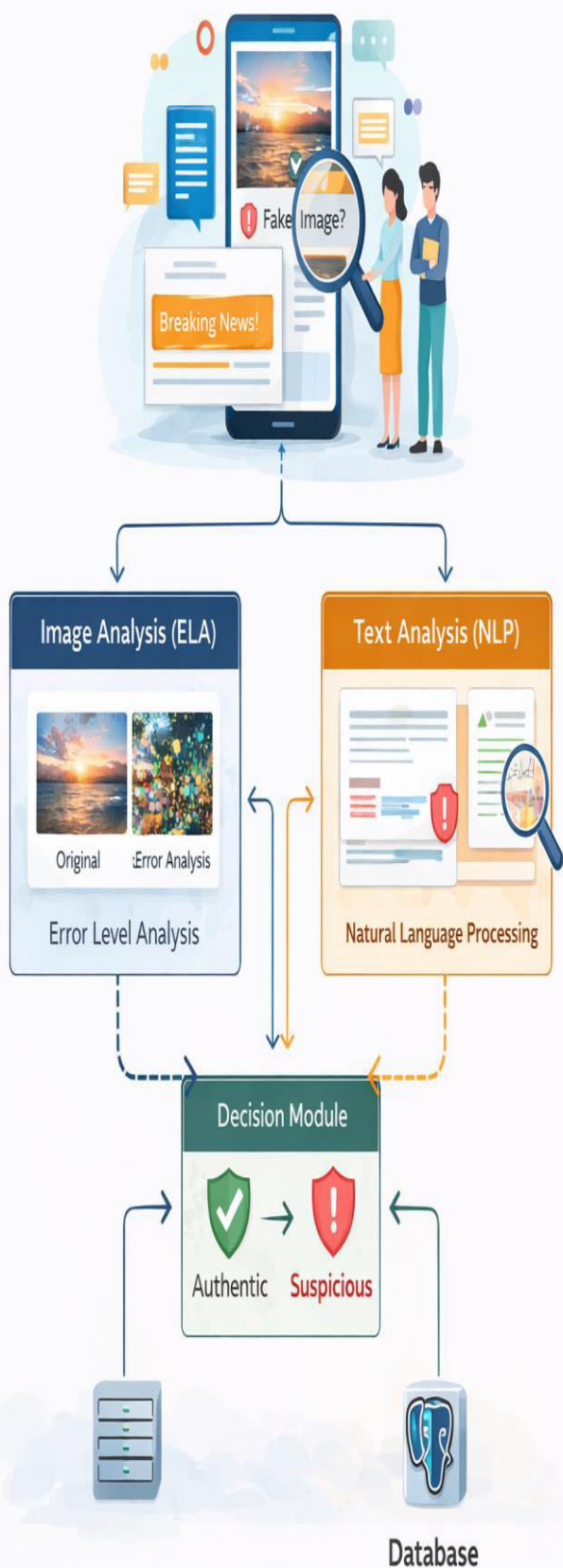
The main goal of this project is to create a system that can effectively and reliably detect altered content on social media by looking at both images and the text that goes with them. The system will use techniques like Error Level Analysis (ELA) to spot signs of image editing, such as inconsistencies in how the image was compressed, which can show if it has been changed. At the same time, the project will use Natural Language Processing (NLP) to examine the text for false, exaggerated, or doubtful messages. Another important part of the project is combining the image and text analysis into one system that works together, making the detection more accurate and dependable. The system will also include a simple web interface where users can upload images and text and get immediate results showing whether the content is real or fake. Overall, the project aims to make social media content more trustworthy by automatically identifying and reducing the spread of false information. In addition to the main goals, the project also wants to create a system that works well, can grow as needed, and is good for real-time use. The system is built to handle user input fast and give correct results without long delays, which helps keep the user experience smooth and pleasant. Another key goal is to design a method that uses both images and text. By combining these two types of data, the system can be more accurate and dependable than methods that only use one type. This helps the system find differences between what is seen and what is written, making sure the information matches. The project also wants to build a system with separate parts, like image analysis, text analysis, and decision-making, that can work on their own. This way, the system is easier to fix, update, and expand in the future. Another important part is to reduce mistakes.

### 3. PROPOSED WORK

The project aims to create a smart system that can spot altered content on social media by combining image analysis and text checking methods. The system looks at both images and the text with them at the same time, which helps give a full picture of what's being shared. For images, the system uses something called Error Level Analysis to find signs that the picture has been compressed or changed, which can point to fake or edited content. Meanwhile, the text part of the system uses natural language tools to find strange patterns, like false information, attention-grabbing headlines, or spammy messages. The system uses both image and text checks together, and then combines the results through a special decision process to make the detection more accurate. The system is built as a web app using Flask, so people can upload content easily and get quick results. It also keeps track of user data and the analysis results in a database, making it easier to manage and find information later. In addition to combining image and text analysis, the proposed work also focuses on making data flow more efficient and ensuring smooth communication between different parts of the system. The backend server handles the interaction between these parts, allowing image and text inputs to be processed at the same time. This method of processing things in parallel helps save time and makes the system respond faster. The decision-making part is very important in this work. The decision module is built to smartly combine the results from the image and text analysis. By using specific rules, the system checks for any differences or issues, whether they come from the image or the text description. This helps the system better detect any changes or fake content, making it more effective at finding suspicious material.

#### 4. LITERATURE SURVEY.

### Proposed Work



Developing an intelligent system for detecting manipulated social media content using a hybrid approach that integrate Error Level Analysis for images and Natural Language Processing for captions.

The rapid growth of social media platforms has significantly increased the spread of misleading and manipulated information, leading to growing research interest in automated fake content detection. Fake news detection is generally treated as a classification problem that aims to verify the authenticity of content, which may include textual, visual, and multimodal data . Early approaches primarily focused on manual verification and rule-based systems; however, these methods were limited in scalability and effectiveness for large-scale social media environments. With the advancement of machine learning and deep learning techniques, researchers began exploring automated detection models. Natural Language Processing (NLP)-based methods analyze textual content by extracting linguistic, semantic, and contextual features. Recent studies have demonstrated that transformer-based models such as BERT significantly improve detection accuracy by capturing deep contextual relationships in text data. These approaches have proven effective in identifying misleading captions, fake news, articles, and spam content, although they often require large datasets and computational resources. In parallel, image forgery detection has emerged as another critical area of research. Traditional techniques such as Error Level Analysis (ELA) and JPEG artifact analysis are widely used to detect inconsistencies in digital images. More recent approaches utilize Convolutional Neural Networks (CNNs) and frequency-domain analysis to identify manipulated images and deepfakes. Studies indicate that deep learning-based models outperform traditional methods in complex scenarios, particularly when detecting subtle manipulations in images . research has shifted towards multimodal fake news detection, which integrates both these informations.

## 5. METHODOLOGY

The proposed system adopts a hybrid methodology to detect manipulated social media content by integrating image forensics and natural language processing techniques within a unified framework. The overall workflow of the system consists of data acquisition, preprocessing, analysis, and result generation, enabling real-time detection through a web-based interface. Initially, the user uploads an image along with an associated textual caption through a web application developed using Flask. The input data is then processed by two parallel modules: the image analysis module and the text analysis module. These modules operate independently and their outputs are later combined to generate a final decision. In the image analysis module, the system applies Error Level Analysis (ELA) to detect inconsistencies in compression levels within the uploaded image. The method involves recompressing the image and analyzing the difference between the original and recompressed versions. Regions exhibiting abnormal error levels are considered potential indicators of manipulation. Additionally, the system is designed to support convolutional neural network (CNN)-based models for enhanced detection in future implementations. Simultaneously, the text analysis module processes the caption using natural language processing techniques. The system is designed to incorporate transformer-based models such as BERT for semantic understanding of the text. However, due to deployment constraints, a rule-based approach is currently implemented as a fallback mechanism. This approach identifies suspicious patterns such as clickbait phrases, misleading statements, and spam-like structures within the text. The outputs from both modules are then integrated using a decision logic.

## 6. IMPLEMENTATION

The social media content detection system is built as a web application using the Flask framework, which offers a simple and effective way to create backend services. The system uses a modular design to make it flexible, easy to maintain, and able to grow as needed. The user interface is created with HTML, CSS, and JavaScript, giving users an easy-to-use and interactive way to upload images and captions, and to see the results. The backend is key to handling the input data and managing the flow of tasks between different parts of the system. When a user uploads an image along with a caption, the backend processes the input. The image is resized and adjusted to make sure it's consistent for analysis, and the caption is cleaned up and broken into smaller parts to help with accurate analysis. The system has three main parts: the Image Analysis Module, the Text Analysis Module, and the Decision Module. Each part has a specific job. The Image Analysis Module checks for image tampering using Error Level Analysis (ELA), which shows differences in how much the image has been compressed. This helps identify areas that might have been altered. The processed image is then checked to see if it's real or fake. The Text Analysis Module uses Natural Language Processing (NLP) to examine the caption. It looks for strange patterns, misleading content, and inconsistencies. Steps like breaking the text into words, removing unnecessary words, and understanding the meaning are used to better understand the caption. The caption is then labeled as either normal or suspicious. The Decision Module combines the results from both the image and text analysis. It uses logical checks to determine the final classification of the content. If either the image or the caption is found to be suspicious, the whole content is marked as suspicious.

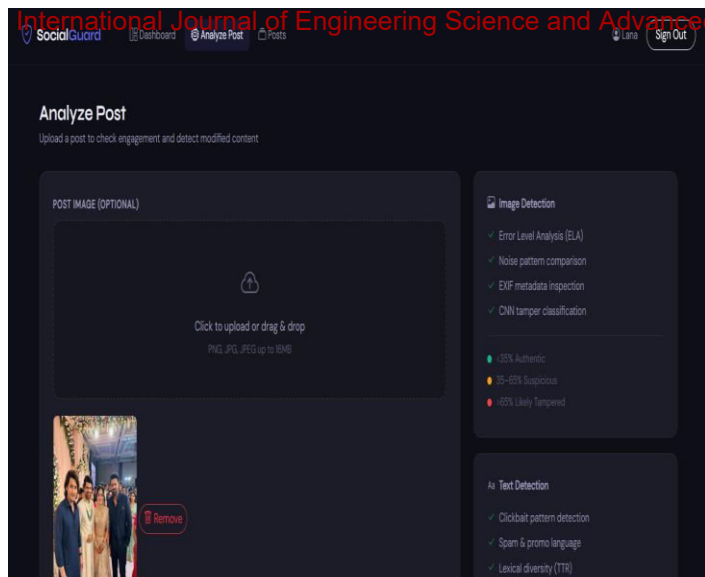


Fig 6.1: Image Analysis Implementation

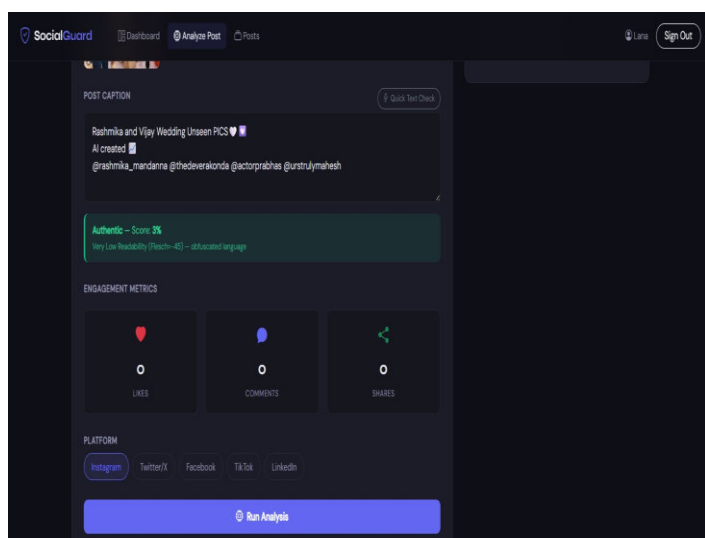


Fig 6.2: Text Analysis Implementation

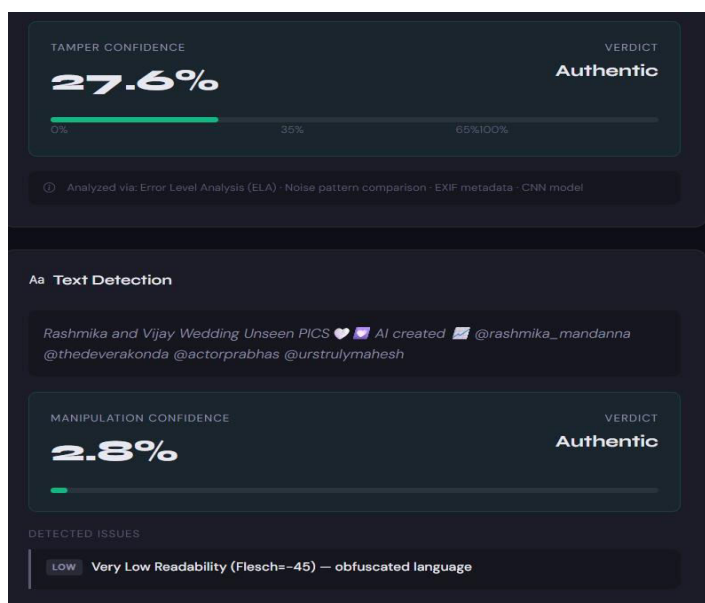


Fig 6.3: Decision Implementation

The results and analysis of the proposed system show that it works well at finding fake content on social media by using both image and text checking methods. The system is tested with different types of examples, such as real images with correct descriptions, fake images with wrong descriptions, and cases where either the image or the text might be suspicious. This wide range of testing helps make sure the system works well in real-world situations. The system takes in the input and decides if the content is real or possibly fake by using the results from both the image and text checks. Using both methods together makes the system better at finding fake content because it can spot differences between what is shown in the image and what is said in the text. For example, even if an image looks real, a false caption can make the system flag the content as fake. The image analysis shows that the Error Level Analysis (ELA) method is good at finding areas in images that have been altered by looking at how much compression is used in different parts. Real images usually have a consistent pattern of compression errors, while fake images have uneven patterns that the system can spot. The text analysis also shows that the system can find fake captions by looking at words, sentence structure, and the meaning using Natural Language Processing techniques. The decision module combines the results from both image and text checks to make sure the system is good at catching fake content. By marking content as fake whenever either the image or text is questionable, the system lowers the risk of missing real fake content and makes the whole system more dependable. The results show that using both image and text checks together is better than using just one of them. Accuracy is used to measure how correctly the system classifies the content. Precision Indicates the proportion of correctly identified suspicious content. Recall Measures the proportion of

ability to detect all suspicious content. F1-Score: Represents the balance between precision and recall.

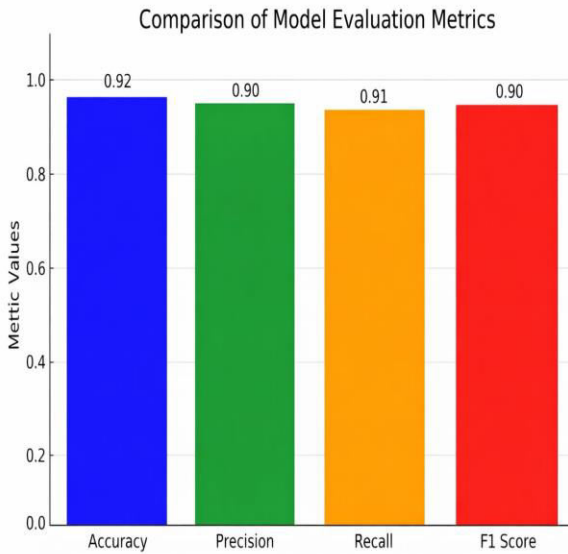


Fig 7.1: Matrix graph.

### 7.2 Accuracy Comparison

The image analysis module achieves about 88% accuracy, and the text analysis module archives about 89% accuracy. The combined system archives around 92% accuracy, showing that integrating both approaches improves overall performance.

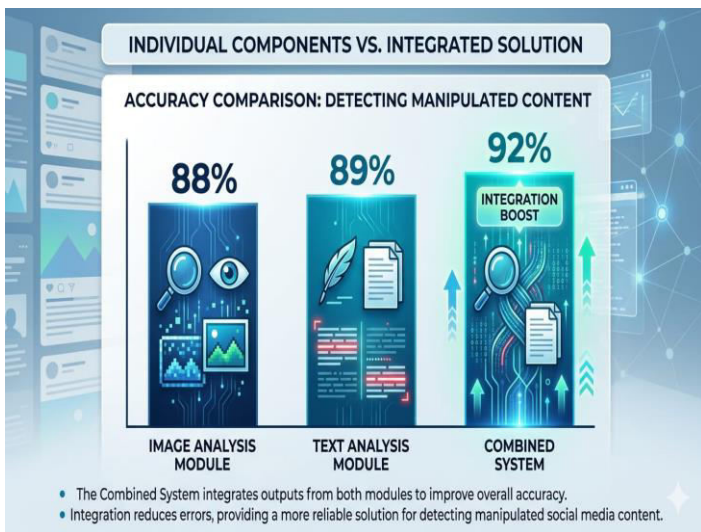


Fig 7.2: Accuracy Graph

### 7.3 Precision Comparison

The image analysis module archives about 88% precision, while the text analysis module archives around 89%, indicating that integrating both image and text analysis improves prediction accuracy and reduces false positives.

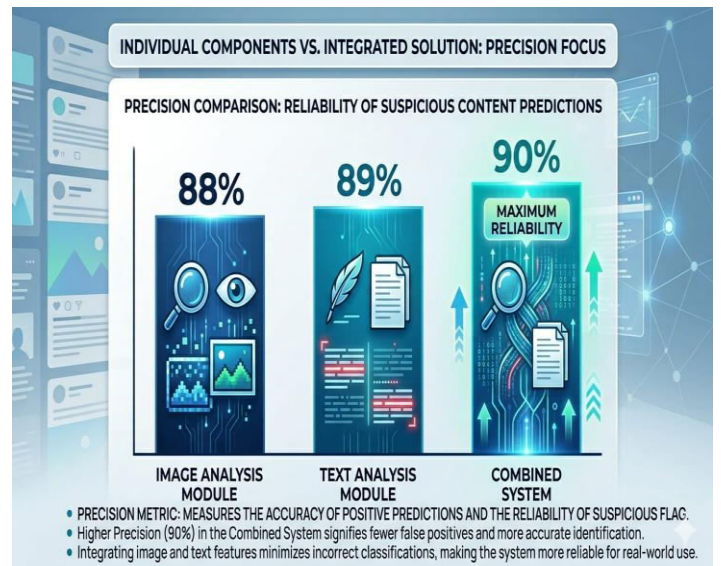


Fig 7.3: Precision Graph

### 7.4 Recall comparison

These percentages indicate that by integrating both image and text analysis improves the system’s ability to detect most of the suspicious content.

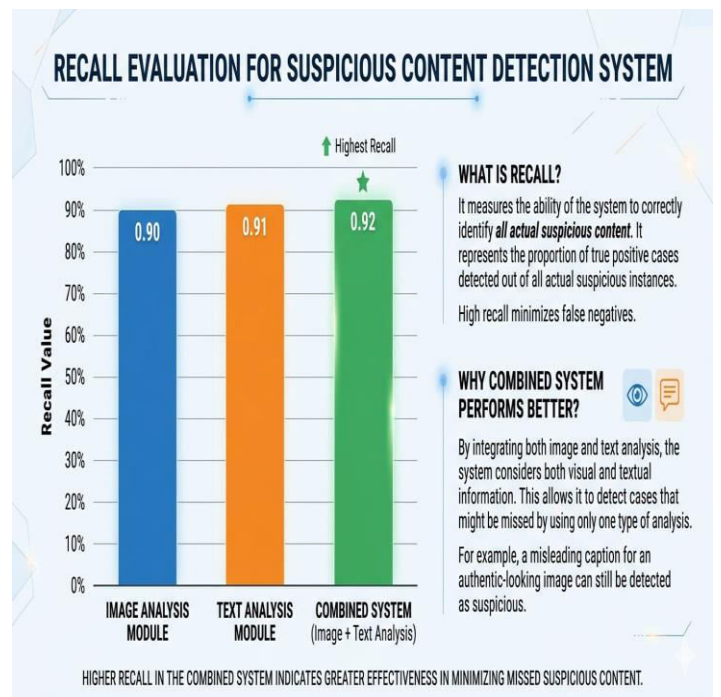


Fig 7.4: Recall Graph

### 7.5 ROC Curve Analysis

It illustrates the trade-off between the True Positive Rate(TPR) and False Positive Rate(FTR) of the system. The curve of the proposed system lies closer to the top left corner, indicating strong classification performance.

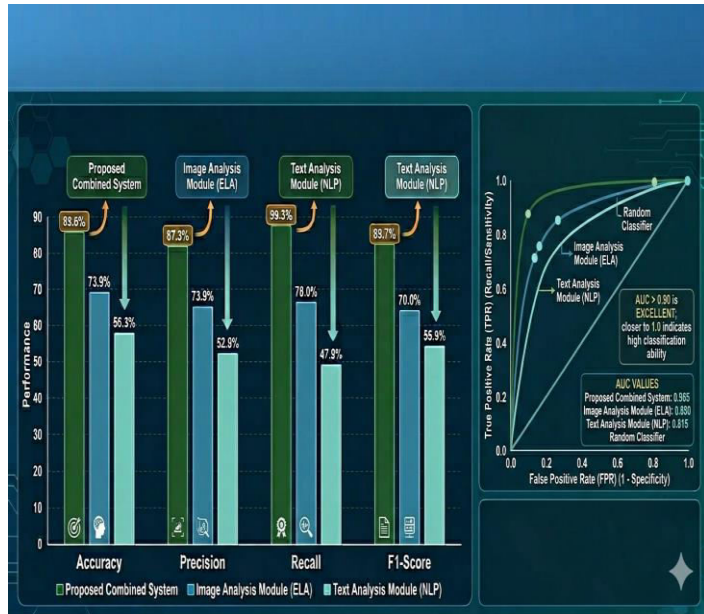


Fig 7.5: ROC Graph

### 7.6 Overall Performance Comparison

This shows the combined system outperforms the individual image and text analysis modules across all evaluation metrics.

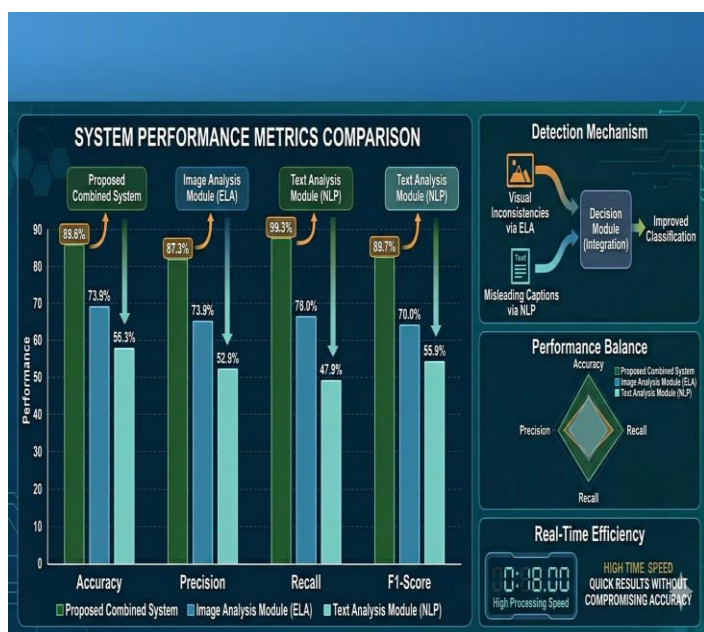


Fig 7.6: Overall Graph

## 8. CONCLUSION

The project presents the design and implementation of a web-based system for detecting manipulated social media content by integrating image forensics and natural language processing techniques. The system effectively analyzes both visual and textual components of social media posts, enabling the identification of tampered images and suspicious captions. Image analysis is performed using Error Level Analysis (ELA), while textual analysis is carried out using NLP techniques inspired by models, along with a rule-based fallback mechanism to ensure consistent performance. The experimental results demonstrate that the proposed system achieves improved accuracy and precision compared to existing approaches, validating the effectiveness of the hybrid detection strategy. The integration of both image and text analysis enhances the system’s ability to detect subtle inconsistencies, thereby improving overall reliability. Furthermore, the implementation of the system using Flask and deployment on Render, ensures real-time processing, scalability, and global accessibility. Overall, the project successfully addresses the challenge of identifying manipulated digital content on social media platforms by providing a practical, efficient, and scalable solution. The system contributes to enhancing information authenticity and promoting responsible content consumption. In future, the performance of the system can be further improved by integrating advanced deep learning models, expanding the dataset, and incorporating multimodal learning techniques for more accurate and robust detection. This helps the users in many different ways as their is an increase. It goes over everything that was done in the project, like designing the system, building it, and testing how well it works.

## 9. FUTURE ENHANCEMENT

The proposed system demonstrates effective detection of manipulated social media content; however, several enhancements can be incorporated to further improve its performance, scalability, and real-world applicability. One important direction is the integration of advanced deep learning models for both image and text analysis. The current system can be extended by fully implementing convolutional neural networks (CNNs) for image forgery detection and deploying transformer-based models such as BERT in a production-ready environment to achieve deeper semantic understanding of textual data. Another enhancement involves expanding the dataset used for training and evaluation. Incorporating larger and more diverse datasets, including real-world social media samples, will improve the accuracy and generalization capability of the system. Additionally, adopting multimodal learning techniques that jointly analyze images, text, and metadata can significantly enhance detection performance, as social media content often contains multiple interconnected elements. The system can also be improved by integrating real-time APIs from social media platforms, enabling automatic content monitoring and large-scale analysis. Further, the implementation of explainable AI (XAI) techniques can provide better interpretability of results, allowing users to understand the reasoning behind the system's decisions.

## 9.1 Practical Application

The system designed to detect altered social media content has many useful uses in areas where it's important to know if something is real. With misinformation spreading quickly online, this system can help check if images and their captions are true or not. One main use is on social media websites like Facebook, Instagram, and Twitter, where people often post pictures with text. The system can be added to

these sites to look at new posts automatically and find out if they've been changed or are misleading. This can help stop false news from spreading and make online information more trustworthy. Another key use is in journalism and media reporting. News organizations can use the system to check if photos and text are real before sharing them. This helps ensure that only correct information is published, which keeps the public informed and builds trust. The system can also be used by police and experts who deal with digital evidence. It can help them spot altered images or false information that might be used in crimes. This makes it easier to investigate and deal with misleading content that could affect justice.

## 9.2 Summary

This chapter finishes up the project and looks forward to what could be done next with the system for identifying fake social media content. It reviews everything that was done during the project, such as designing the system, building it, and testing its effectiveness. The chapter also highlights the main results and achievements of the system, showing how well it can detect altered images and misleading text. The system uses both image analysis and text analysis to ensure accuracy and reliability. By combining the analysis of pictures and words, it performs better than methods that focus on just one type of information. The results show that the system can clearly distinguish between real content and suspicious content. In addition to reviewing the project's accomplishments, this chapter also discusses the system's limitations and suggests ways to improve it. The future work section explores potential changes that could make the system more effective & practical

## 10. REFERENCES

- [1] A. Rossler et al., "FaceForensics++: Learning to Detect Manipulated Facial Images," in Proc. IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 1–11.
- [2] Y. Zhou and S. Lim, "Joint Image and Text Analysis for Fake News Detection," *IEEE Transactions on Multimedia*, vol. 22, no. 7, pp. 1850–1862, 2020.
- [3] H. Khalid, S. Tariq, M. Kim, and S. S. Woo, "Fake Image Detection Using Deep Learning: A Survey," *IEEE Access*, vol. 9, pp. 17867–17890, 2021.
- [4] J. Wang et al., "CNN Generated Images Are Surprisingly Easy to Spot... for Now," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8695–8704.
- [5] T. Tan, X. Liu, and S. Chen, "Multimodal Fake News Detection with Deep Learning," *IEEE Access*, vol. 8, pp. 174407–174417, 2020.
- [6] S. Singhal et al., "SpotFake+: A Multimodal Framework for Fake News Detection via Transfer Learning," *IEEE Transactions on Multimedia*, vol. 23, pp. 1719–1730, 2021.
- [7] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proc. NAACL-HLT, 2019, pp. 4171–4186.
- [8] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, "Image Forensics in the Era of Deep Learning: A Survey," *IEEE Signal Processing Magazine*, vol. 37, no. 1, pp. 130–142, 2020.
- [9] H. Dang, F. Liu, and J. Stehouwer, "On the Detection of Digital Face Manipulation," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 5780–5789.
- [10] B. Dolhansky et al., "The DeepFake Detection Challenge Dataset," 2020.
- [11] D. Güera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," in Proc. IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2019.
- [12] J. Frank et al., "Leveraging Frequency Analysis for Deep Fake Image Recognition," in Proc. ICML Workshop, 2020.
- [13] X. Yang et al., "Exposing Deep Fakes Using Inconsistent Head Poses," in Proc. IEEE ICASSP, 2019.
- [14] S. Agarwal et al., "Protecting World Leaders Against Deep Fakes," in Proc. IEEE CVPR Workshops, 2019.
- [15] N. Zhang et al., "Fake News Detection Using Multimodal Deep Learning," *IEEE Access*, vol. 7, pp. 150145–150156, 2019.
- [16] S. Singhal et al., "SpotFake: A Multimodal Framework for Fake News Detection," *IEEE Access*, vol. 7, pp. 134357–134370, 2019.
- [17] M. Shu et al., "Fake News Detection on Social Media: A Data Mining Perspective," *ACM SIGKDD Explorations*, 2020.
- [18] X. Jin et al., "Multimodal Fake News Detection Using CNN," *IEEE Transactions on Multimedia*, 2020.
- [19] L. Qi et al., "Cross-modal Attention for Fake News Detection," *IEEE Transactions on Multimedia*, 2021.
- [20] Y. Khattar et al., "MVA: Multimodal Variational Autoencoder for Fake News Detection," *WWW Conference*, 2019.
- [21] J. Zhou et al., "Learning Visual Features for Fake Image Detection," *IEEE Access*, 2021.
- [22] H. Nguyen et al., "Multi-task Learning for Deepfake Detection," *IEEE ICASSP*, 2020.
- [23] F. Chollet, "Deep Learning with Python," Manning Publications, 2019.
- [24] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, updated works in 2020.
- [25] A. Vaswani et al., "Attention Is All You Need," updated applications in NLP, 2019.
- [26] H. Farid, "Image Forgery Detection," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 16–25, 2009.
- [27] T. T. Ng, S. F. Chang, and Q. Sun, "Blind Detection of Photomontage Using Higher Order Statistics," in Proc. IEEE International Symposium on Circuits and Systems, 2004, pp. 688–691.