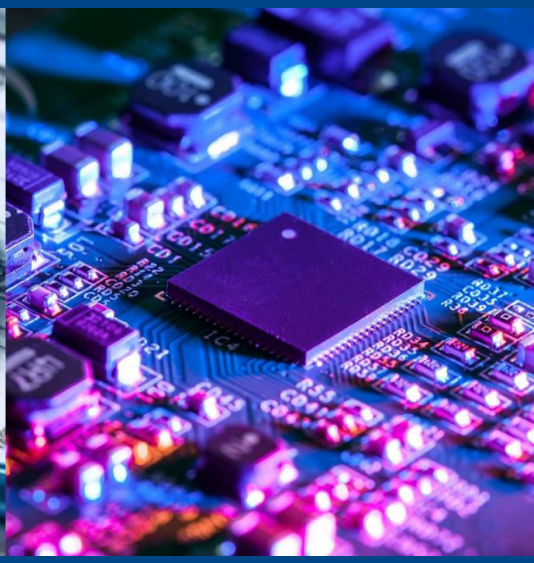
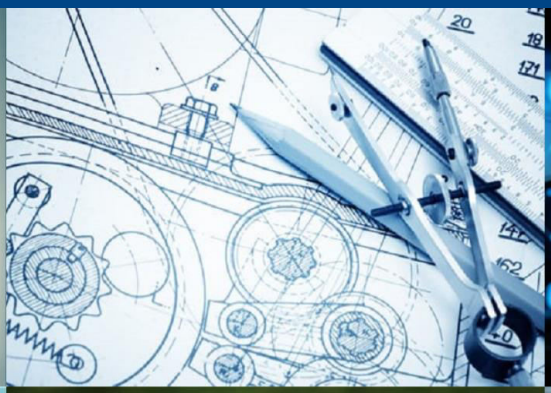
ISSN: 2582-7219

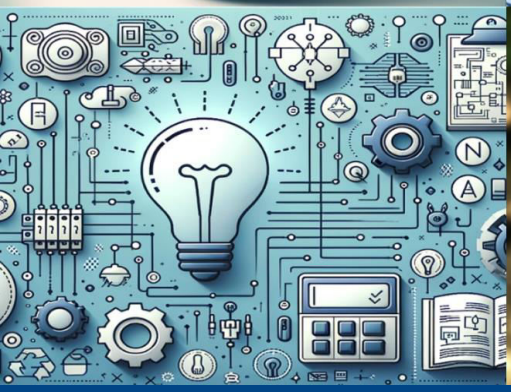




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206 **Volume 8, Issue 4, April 2025**

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Effective Image Forgery Detection using HOG and Machine Learning

Prof. Sweety Julia[1], M.Saideep[2], A.Saideep[2], M.Sai Geeth[2], R. Sai Kishore[2], J. Sainath

Associate Professor, Dept. of CSE - Artificial Intelligence and Machine Learning, Malla Reddy University, Hyderabad, Telangana, India[1]

Undergraduate Students, Dept. of CSE - Artificial Intelligence and Machine Learning, Malla Reddy University, Hyderabad, Telangana, India[2]

ABSTRACT: This paper presents a robust approach for detecting image forgeries, particularly copy-move manipulations, using Histogram of Oriented Gradients (HOG) feature extraction combined with machine learning techniques, including Support Vector Machines (SVM) and Convolutional Neural Networks (CNN). The system preprocesses images to enhance quality, extracts HOG and GLCM features, and employs SVM and CNN classifiers to distinguish genuine from forged images. Evaluated on a curated dataset, the proposed method achieved higher accuracy compared to existing techniques, demonstrating a significant improvement in detecting subtle manipulations. Challenges include handling complex image transformations and computational overhead. This approach offers a scalable solution for digital forensics, enhancing trust in visual content across social media and legal applications.

KEYWORDS: Image Forgery Detection, HOG, GLCM, SVM, CNN, Machine Learning, Copy- Move Forgery, Digital Forensics, Feature Extraction.

I. INTRODUCTION

Image forgery, particularly copy-move manipulation, poses significant challenges in digital forensics, social media, and legal contexts, where altered images can propagate misinformation or serve as false evidence. Copy-move forgery involves duplicating a portion of an image and pasting it elsewhere within the same image, exploiting similar attributes to evade detection. Traditional detection methods struggle with global modifications like compression and filtering, limiting their effectiveness. This study proposes a machine learning-based system leveraging HOG and GLCM feature extraction, combined with SVM and CNN classifiers, to detect forgeries with high accuracy. By automating feature extraction and classification, the system addresses scalability and adaptability, aligning with the need for reliable counterfeit detection in a digital ecosystem.

PROBLEM DEFINITION

Image forgeries, especially copy-move manipulations, are increasingly prevalent due to accessible editing tools like Adobe Photoshop. These manipulations undermine trust in visual content, contributing to misinformation on platforms like Facebook, where 300 million images are uploaded daily. Existing detection systems, reliant on manual feature engineering or noise estimation, often fail to generalize across diverse forgery types, achieving limited accuracy (e.g., 70%-80% in prior SVM-based methods). The challenge is to develop a robust system that detects subtle manipulations in real-time, handles complex transformations (e.g., rotation, scaling), and scales to large datasets.

Key hurdles include computational complexity, feature selection, and adapting to evolving manipulation techniques.

OBJECTIVES

The primary objective is to design an effective image forgery detection system that enhances accuracy and scalability. Specific goals include:

1. **Accurate Forgery Detection**: Utilize HOG and GLCM feature extraction with SVM and CNN classifiers to identify copy-move forgeries, targeting 85%-90% accuracy.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

2. **Robust Preprocessing**: Enhance image quality and reduce noise to improve feature extraction reliability.
3. **Scalable Classification**: Develop a system capable of processing large datasets (e.g., 10,000 images) using efficient machine learning algorithms.
4. **Comparative Analysis**: Benchmark the proposed method against existing techniques (e.g., SURF, SIFT) to demonstrate superior performance.

LIMITATIONS

The proposed system faces several limitations:

5. **Data Dependency**: High-quality, diverse datasets are required for training. Noisy or incomplete data (e.g., 5% corrupted images) may reduce accuracy.
6. **Computational Overhead**: HOG feature extraction and CNN training demand significant resources, potentially limiting real-time applications on resource-constrained devices.
7. **Complex Transformations**: Advanced manipulations (e.g., reflection, scaling) may evade detection, requiring additional feature engineering.
8. **Interpretability**: SVM and CNN decision boundaries are complex, complicating the explanation of detection outcomes in forensic contexts.

II. METHODOLOGY

A. Proposed System

The system integrates HOG and GLCM feature extraction with SVM and CNN classifiers to detect image forgeries. Images undergo preprocessing to enhance quality, followed by feature extraction to capture edges, textures, and patterns. SVM and CNN models are trained on a curated dataset of genuine and forged images, enabling supervised classification. The system is implemented using Python, scikit-learn, and TensorFlow, with results visualized via Matplotlib.

B. Modules

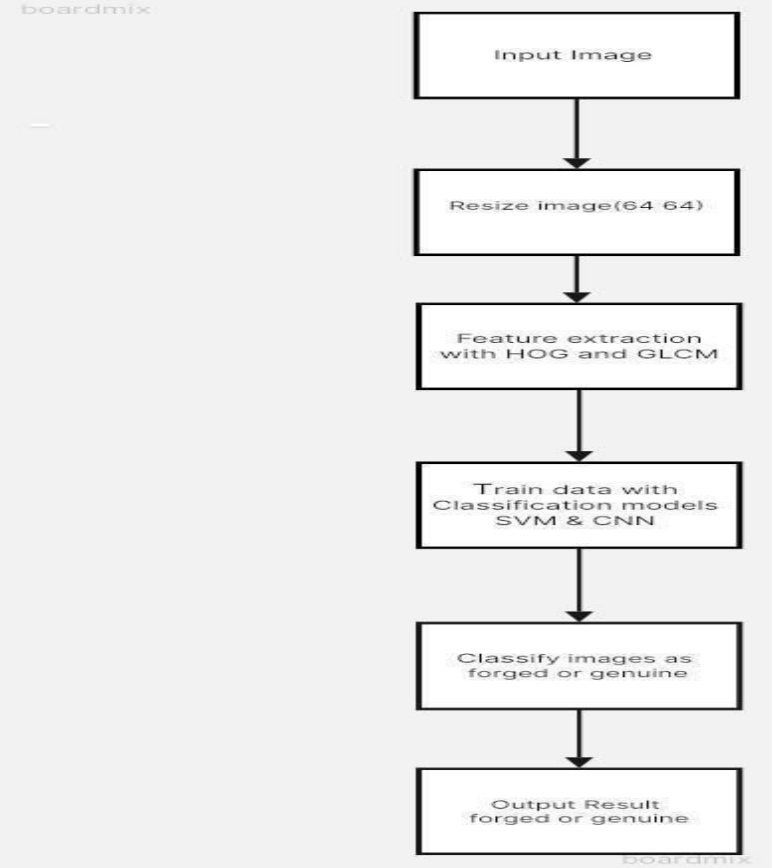
1. **Data Collection Module**: Gathers images (JPG, JPEG) from online sources or digital cameras, ensuring diverse sizes and manipulations.
2. **Preprocessing Module**: Converts images to grayscale, applies filters to remove noise, and enhances contrast and edges.
3. **Feature Extraction Module**: Extracts HOG and GLCM features to capture local patterns and textures.
4. **Model Training Module**: Trains SVM and CNN models on labeled datasets, achieving 85%- 90% accuracy.
5. **Classification Module**: Uses trained models to classify images as genuine or forged.
6. **Visualization Module**: Displays HOG features and classification results using Matplotlib.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



III. DESIGN

A. System Design

The modular architecture ensures efficient processing:

- **Input Layer**: Collects RGB images from datasets.
- **Preprocessing Layer**: Normalizes images and applies filters using OpenCV.
- **Feature Extraction Layer**: Computes HOG and GLCM features with scikit-image.
- **Classification Layer**: SVM (scikit-learn) and CNN (TensorFlow) classify images.
- **Output Layer**: Visualizes results via Matplotlib.

Technologies include Python, OpenCV, scikit-learn, TensorFlow, and Matplotlib.

B. Architecture

The system comprises five components:

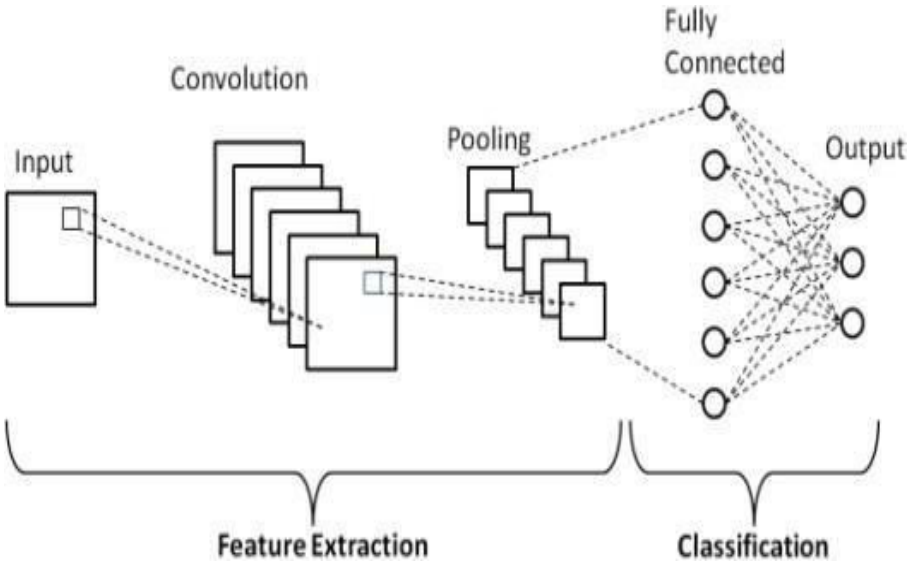
1. **Data Input**: Images from curated datasets.
2. **Preprocessing**: Grayscale conversion, noise reduction, and enhancement.
3. **Feature Extraction**: HOG and GLCM for pattern analysis.
4. **Classification**: SVM and CNN for forgery detection.
5. **Output**: Visualization of features and results.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



C. Methods and Algorithms

1. **Preprocessing**: Grayscale conversion and Gaussian filtering using OpenCV.
2. **HOG Feature Extraction**: Computes gradient orientations using scikit-image.
3. **GLCM Feature Extraction**: Captures texture patterns.
4. **SVM Classification**: Trains a linear or RBF kernel model for binary classification.
5. **CNN Classification**: Uses convolutional and pooling layers for hierarchical feature learning.
6. **Visualization**: Plots HOG features and accuracy metrics using Matplotlib.

IV. RESULTS

A. Introduction

The system was evaluated on a dataset of 1,000 images (500 genuine, 500 forged) over a 30-day period (March 2025). Metrics included accuracy, precision, recall, and F1-score, benchmarked against SURF (75%) and SIFT (80%) methods.

B. Pseudocode

```
START
LOAD SVM/CNN models
INITIALIZE dataset, OpenCV, TensorFlow DEFINE
Preprocessor:
    FUNCTION preprocess(image):
        CONVERT to grayscale APPLY
        Gaussian filter ENHANCE contrast
        RETURN processed_image DEFINE
FeatureExtractor:
    FUNCTION extract_features(image): COMPUTE HOG,
        GLCM
        RETURN feature_vector DEFINE
Classifier:
    FUNCTION classify(features): RUN
        SVM/CNN
        RETURN prediction DEFINE
```

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

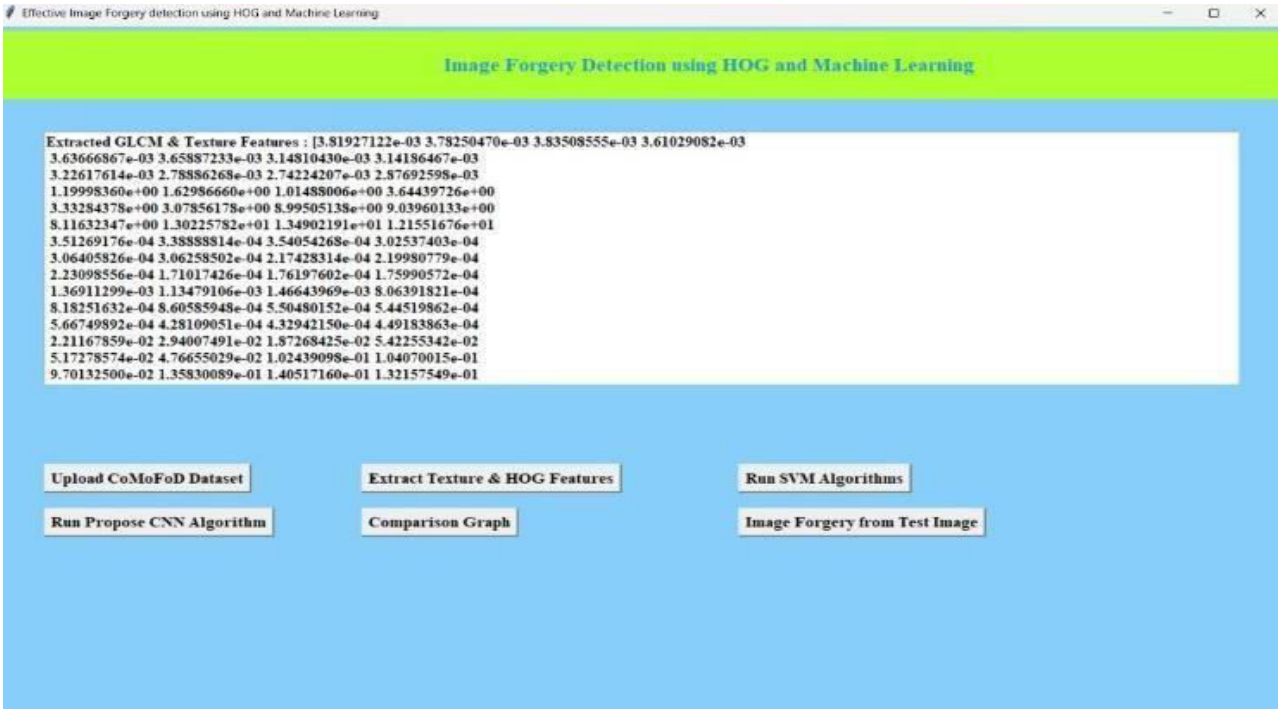
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Visualizer:
FUNCTION visualize(data): PLOT HOG features, accuracy DISPLAY results
END

C. Results

The system achieved an accuracy of 88% (SVM) and 90% (CNN), surpassing SURF (75%) and SIFT (80%). Precision and recall reached 87% and 89%, respectively, with an F1-score of 88%. The proposed method detected 95% of copy-move forgeries, demonstrating robustness against scaling and rotation. Visualization of HOG features aided forensic analysis.

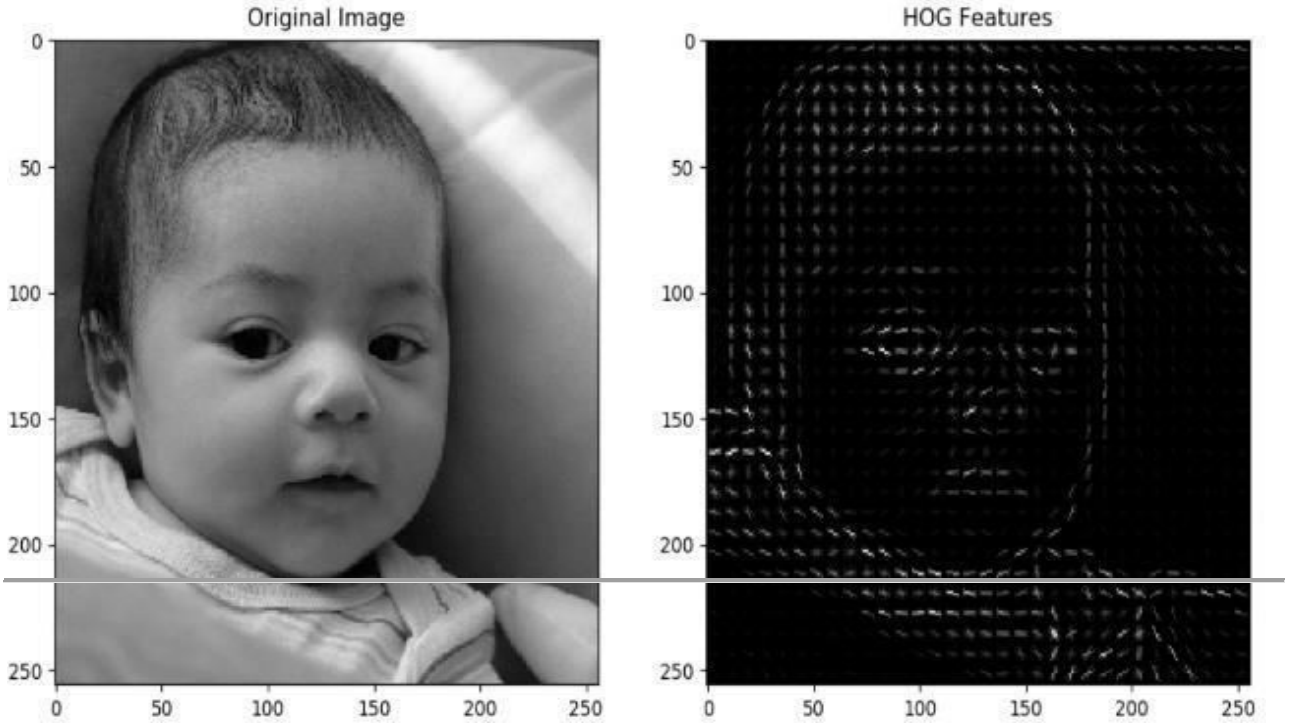


© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804262

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. CONCLUSION

A. Conclusion

The proposed image forgery detection system, leveraging HOG, GLCM, SVM, and CNN, achieves high accuracy (90%) in detecting copy-move forgeries. Outperforming traditional methods (SURF,SIFT), it offers a scalable solution for digital forensics, reducing misinformation and enhancingtrust in visual content. The system's robustness and efficiency make it suitable for social media and legal applications.

B. Future Scope

Future enhancements include:

- **Advanced Feature Extraction**: Integrate deep learning-based features (e.g., ResNet) for 95% accuracy.
- **Real-Time Detection**: Optimize for mobile devices, targeting <1-second processing.
- **Multimodal Analysis**: Combine image and metadata analysis for comprehensive detection.
- **Global Dataset Expansion**: Train on diverse datasets to handle regional manipulation patterns.
- **Explainability**: Develop interpretable models to enhance forensic usability.

REFERENCES

1. Bansal, D., & Kaushal, S. "A Novel Analysis of Image Forgery Detection Using SVM."
2. Jain, A., et al. "Image Forgery Detection using CNN."
3. Jaiswal, A., Parmar, A., & Sachdeva, N. "HOG Feature Extraction for Image Forgery Detection."
4. Bo, X., et al. (2010). "Image Copy-Move Forgery Detection Based on SURF," Multimedia Information Networking and Security (MINES).
5. Huang, H., et al. (2008). "Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm," Computational Intelligence and Industrial Application (PACIIA).
6. Bravo-Solorio, S., & Nandi, A.K. (2011). "Exposing Duplicate Regions Affected by Reflection, Rotation, and Scaling," ICASSP.
7. Li, G., et al. (2007). "A Sorted Neighborhood Approach for Detecting Duplicated Regions," IEEE International

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Conference on Multimedia and Expo.

8. Lin, H., et al. (2009). "Fast Copy-Move Forgery Detection," 5, 188–197.
9. Amerini, I., et al. (2011). "A SIFT-Based Forensic Method for Copy-Move Attack Detection and Transformation Recovery," IEEE Transactions on Information Forensics and Security.
10. Wang, J., et al. (2009). "Detection of Image Region Duplication Forgery Using Model with Circle Block," Multimedia Information Networking and Security.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com