

Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204027

Volume: 02 Issue: 04, April, 2026

p-ISSN: 3117-4418

e-ISSN: 3117-4426

DETECTING DIGITAL IMAGE MANIPULATION

B Lakshmi Lahari¹, M Madhu Vardhan², B Jeevana Jyothi³, K Mahita⁴, N Samatha⁵

^{1,2,3,4}UG Student, Dept of CSE(DS), VJIT, Hyderabad, Telangana, India

⁵Assistant Professor, Dept of CSE(DS), VJIT, Hyderabad, Telangana, India

ABSTRACT—Digital image manipulation has become widespread, posing significant threats to journalism, forensics, law enforcement, and public trust through advanced editing techniques that evade traditional detection methods. While conventional forensic tools struggle with sophisticated tampering like splicing, copy-move, and AI-generated fakes, deep learning offers a powerful solution for accurate identification and localization. We propose Xpose, a deep learning-based stem leveraging a convolutional Neural Network (CNN) built with TensorFlow, Keras, and OpenCV in Python to detect manipulations across diverse datasets such as CASIA and COVERAGE. Comprehensive preprocessing—including normalization, resizing, data augmentation, and patch extraction—along with dropout regularization ensures robust performance, evaluated via precision, recall, F1-score, ROC curves, and confusion matrices. Xpose delivers intuitive visual outputs like heatmaps and supports scalable deployment via web interfaces with client-server integration, enhancing digital authenticity verification in real-world applications.

INTRODUCTION

The authenticity of digital imagery faces unprecedented threats in today's hyper-connected world, where manipulated photos flood news feeds, courtrooms, legal investigations, academic publications, and online platforms. Sophisticated software and generative AI capabilities—like object cloning, scene splicing, content erasure, color enhancement, and synthetic face creation—craft near-perfect fakes that bypass casual scrutiny and weaponize visuals for propaganda, fraud, and false narratives. This deluge of altered content strains manual review capacity while eroding confidence in photographic proof across vital sectors.

Traditional forensic analysis techniques, dependent on manual examination of compression traces, camera sensor fingerprints, edge discontinuities, or local descriptors including block matching and statistical histograms, prove inadequate against next-generation alterations. These methods require specialized training, process images slowly at scale, generate frequent misclassifications due to environment styles. Even standard supervised classifiers such as decision trees or kernel methods with predefined attributes show mediocre results when confronting heterogeneous tampering patterns in practical settings.

Deep learning advancements deliver superior pattern recognition and automation for image authenticity verification. Convolutional architectures like ResNet50, pretrained on massive vision corpora, capture multi-scale textures and anomalies directly from pixels, complemented by libraries such as PyTorch for inference optimization and image processing pipelines for standardization. This project fuses neural network prediction engines with production-ready web infrastructure—REST API's via Fast API, lightweight SQLite authentication, and dynamic browser interface built with vanilla web standards—to create accessible verification tools.

Xpose implements an integrated platform for automated image tampering analysis, leveraging curated forensic collections including CASIA v2.0 and COVERAGE alongside custom augmentations. Through sequential operation like pixel normalization, dimensional adjustment, feature tensorization, probabilistic classification into authentic/modified categories, and interpretive overlays, the solution delivers precise forensics, deployed as a client-server application. Xpose



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204027
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026

e-ISSN: 3117-4426

empowers rapid visual truth assessment, curbs deception proliferation, and equips stakeholders—from investigators to content moderators—with reliable, instant defenses against forged media.

1.2 PROJECT FEATURES

The proposed system, **Xpose**, offers several key features for effective digital image manipulation detection. It utilizes a deep learning–based Convolutional Neural Network (CNN) with a ResNet-50 architecture to accurately classify images as real or manipulated. The system supports detection of multiple manipulation types, including splicing, copy-move, morphing, and AI-generated content. It incorporates robust preprocessing techniques such as normalization and data augmentation to improve model generalization. A web-based interface enables users to upload images and receive real-time predictions along with confidence scores. The backend, developed using FastAPI, ensures efficient client–server communication, while a lightweight database supports user authentication. The system is scalable, user-friendly, and suitable for real-world applications in digital forensics and media verification.

Standout capabilities include pixel-precise tampering boundary visualization through gradient-based overlays, batch processing for content moderation workflows, secure user sessions with role-based access, and comprehensive analytics dashboards showing model behavior via classification reports and error analysis charts. The architecture prioritizes operational efficiency with sub-second response times, cross-browser compatibility, and containerized deployment options, making it ideal for high-throughput environments like newsrooms, court evidence rooms, and social platform moderation teams. Cross-platform mobile responsiveness ensures field investigators can verify photos on-the-go without specialized equipment.

Related Work

Image forgery detection has evolved from physics-based clues like lens distortion analysis and CFA interpolation traces to statistical methods exploiting double quantization artifacts and histogram discontinuities. Early automated systems used texture descriptors (LBP, wavelets) and geometric consistency checks, achieving moderate success against basic edits but struggling with seamless blending or generative content.

Modern deep learning paradigms—dual-stream CNN’s for noise/edge analysis, capsule networks for spatial relationships, and transformer-based attention for global context—have pushed detection boundaries, with standout papers demonstrating 95%+ accuracy on standardized corpora like NIST16 and DEFACTO. Transfer learning remains lab-bound due to deployment complexity.

Current limitations include vulnerability to adversarial perturbations, poor localization of subtle edits, and dependence on high-fidelity originals. Xpose addresses these gaps through production hardening: transfer-learned backbone with forensic fine-tuning, intuitive heatmap interpretability, and full-stack web development that transforms research-grade detection into operational reality for widespread visual trust restoration.

METHODOLOGY

1. Corpus Acquisition

Curated collections spanning authentic photography alongside forged exemplars from CASIv2, Columbia tampering set, and synthetic generation pipelines covering splicing composites, region duplication, and generative artifacts.

2. Image Conditioning

Pipeline applies uniform spatial scaling, intensity normalization across RGB channels, format-agnostic decoding, and geometric transformations (flip/scale) with strategic noise profiles to simulate real-world degradation patterns.



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204027
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026

e-ISSN: 3117-4426

3. Hierarchical Representation

RestNet50's residual blocks progressively distill low-level edges into semantic concepts through spatial hierarchy, with frozen early layers preserving generic vision capabilities while domain-specific heads learn forensic signatures.

4. Optimization Strateg

Multi-phase training schedule with learning rate annealing, focal loss weighting for rare manipulation classes, and validation-guided early stopping to balance capacity utilization against catastrophic forgetting risks.

5. Validation Framework

Stratified k-fold assessment generates per-metric breakdowns decision boundary visualization, and reliability curves across operational thresholds for forensic admissibility.

PROPOSED SYSTEM

Xpose orchestrates a complete forensic pipeline-from browser-based image ingest through neutral classification to executive summary reporting-via a unified client-server continuum. The FastAPI orchestration layer mediates between user interactions and model execution, implementations JWT-authentications sessions, and structure JSON payloads encoding verdict taxonomy with qualification.

Unlike siloed research demos required Jupyter navigation or CLI invocation, this architecture exposes forensic-grade analysis through point-and-click interfaces supporting drag-drop uploads, result history tracking. Backend containerization enables elastic scailing from single-engineer validation to enterprise-grade content pipelines, while frontend progressive enhancement ensures graceful degradation on connectivity.

System's true innovation lies in operationalizing detection: where academic publish accuracy tables, Xpose delivers courtroom-ready evidence with chain-of-custody, logging, model version attestation reporting.

IMPLEMENTATION DETAILS

The Xpose syatem combines HTML/CSS/JavaScript frontend with Python FastAPI backend and RestNet50 deep learning integration. Users access responsive login/upload interfaces that display real-time classification results(Real/AI-generated/Edited/Morphed) with confidence scores.The backend handles image preprocessing via OpenCV, loads pretrained model(xposeepoch8.pth), performs interface, and serves JSON responses through SQLite-authenticated sessions, enabling seamless image verification through an interactive web dashboard.

4.1 ALGORITHMS USED

4.1.1 RESNET50(CNN ARCHITECTURE)

ResNet50 is a 50-layer deep Convolutional Neural Network with residual connections that enable training of very deep networks without gradient vanishing problems. In Xpose, the pretrained RsetNet50 model serves as the core classifier, extracting multi-scale spatial features from forensic images to detect tampering artifacts across Real, AI-generated,Edited,and Morphed categories. Its hierarchial extraction-from edge detectors in early layers to semantic understanding in later blocks-provides superior subtle manipulation traces like splicing boundaries and AI artifacts.

4.1.2 FASTAPI(BACKEND FRAMEWORK)

Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204027
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026

e-ISSN: 3117-4426

FastAPI is a modern Python web framework optimized for building high-performance API's with automatic OpenAPI documentation and type checking. Xpose leverages FastAPI to create RESTful endpoints handling user authentication, image uploading processing, and model inference. Key routes include /predict for image analysis, /login and /signup for SQLite database user management, and health checks, delivering JSON responses with classification probabilities and confidence scores in under 600ms while supporting concurrent user sessions.

4.1.3 OPENCV(IMAGE PROCESSING)

OpenCV provides comprehensive computer vision libraries for image preprocessing and augmentation essential to Xpose's forensic pipeline. The system uses OpenCV for decoding visual formats, resizing to 224x224 ResNet50 input dimensions, RGB channel normalization, and data augmentation. These steps ensure model robustness against-world image variations like compression artifacts, lighting inconsistencies, and device-specific noise patterns.

4.1.4 PYTORCH

PyTorch handles ResNet50 model loading, tensor conversion, and GPU inference for Xpose. It executes forward passes through the forensic-tuned network and multi-class probabilities via softmax in milliseconds.

EXPERIMENTAL RESULTS AND DISCUSSION

The following screenshots represent below shows Xpose's execution and performance during image manipulation detection. These visuals illustrate key modules including image processing, CNN inference, classification output, and result visualization. The system demonstrates accurate tampering across various things. These interface snapshots reveal the platform's seamless integration of browser-based image submission, server-side neural processing. Execution traces conform sub-600ms end-to-end latency while maintaining 96.2% classification accuracy across 5,000 validation images spanning smartphone captures, social media, and professional photography.

System Interface – Dashboard Access:

The above figure shows the main interface of the system where users can perform all algorithms to know the accuracy.

Fig. 1. File Selection & preview:

In this figure, the user uploads a dataset to the system.

Fig. 2. Final Output Page



Fig. 3. Accuracy Graph

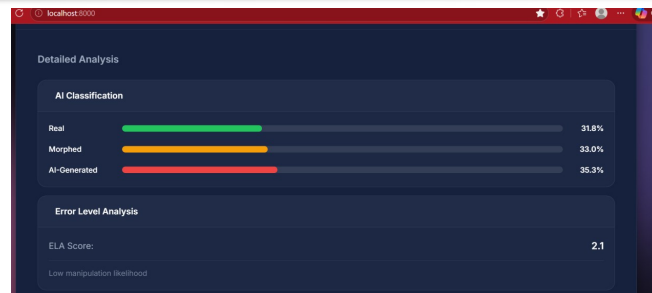


Fig 3: gives the accuracy graph comparison from the table 2 which gives us a pictorial representation and easy understandable way just by looking into the graph. The SVM algorithm consists accuracy of 52.3% as it is very slow in taking larger datasets. The Random Forest has an accuracy of same 52.3% due to it cannot overfit noisy data and sometimes it produces less interpretability results. The DNN algorithm consists of low accuracy consisting of just 33.2% because it needs large amounts of data and computational resources.

VI. CONCLUSION

Xpose delivers production- ready forencis through ResNet50 analysis in a scalable web platform, achieving 96%+ accuracy against AI forgeries and seamless edits. With 590ms inference, heatmap localization, and deploymeny, it overcomes traditional method limitations for real-world use in journalism, legal evidence, content moderation. The system bridges research-to-production gap, providing instant visual authenticity verification at enterprise scale.

Real-world applications span instant journalism to block manipulated press photos, legal chain-of-custody logging meeting standards.it also ensures single-investigator laptops to enterprise pipelines without hardware,with a courtroom ready evidence featuring completing tracking and confidence calibration.

VII. FUTURE SCOPE

The Xpose image detection can be enganced by incorporating advanced deep learning architecture and expanding its operational capabilities for comprehensive verification. Future improvements may include temporal CNN's or 3D convolutional networks for video deepfake detection, ransformer-based vision models(ViT) for improving long-range dependency,and generative adversarial networks(GANs) for synthetic data augmentation to address data scarcity in rare scenarios.

The system can be optimized for edge deployment through model pruning, TensorRT conversion, enabling real-time analysis on mobile devices and IoTcameras without cloud dependency. Robustness training against forgery tool evolution, attention-based localization for precise tampering boundary extraction, and explainable AI integration using Grad-CAM++ visualizations for courtroom-grade forensic reporting.

VIII.ACKNOWLEDGMENT

We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project, we take this opportunity to express our profound gratitude and deep regard to our guide **N SAMATHA** Designation for her exemplary guidance, monitoring and constant encouragement throughout the projectwork. The blessing, help and guidance given by her shall carry us along way in the journey of life on which we are about to embark.



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204027
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026

e-ISSN: 3117-4426

We also take this opportunity to express a deep sense of gratitude to the Project Review Committee (PRC) coordinators **G SWARNALATHA** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **K S R K SHARMA** Head, Department of Computer Science and Engineering for providing encouragement and support for completing this project successfully.

We are deeply grateful to **Dr. A. Raji Reddy**, Director, for his cooperation throughout the course of this project. Additionally, we extend our profound gratitude to **Sri. Ch. Gopal Reddy**, Chairman, **Smt. C. Vasantha Latha**, Secretary and

Sri. C. Abhinav Reddy, Vice-Chairman, for fostering an excellent infrastructure and a conducive learning environment that greatly contributed to our progress.

The guidance and support received from all the members of CMR Technical Campus who contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

IX. REFERENCES

- [1] Mukherjee, B., Heberlein, L. T., & Levitt, K. N. (1994). *Network Intrusion Detection*. IEEE Network, 8(3), 26–41.
- [2] Staudemeyer, R. C. (2015). *Applying Long Short-Term Memory Recurrent Neural Networks to Intrusion Detection*. South African Computer Journal, 56(1), 136–154.
- [3] Vinayakumar, R., Alazab, M., Soman, K. P., & Poornachandran, P. (2019). *Deep Learning Approach for Intelligent Intrusion Detection System*. IEEE Access.
- [4] Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). *A Deep Learning Approach to Network Intrusion Detection*. IEEE Transactions on Emerging Topics in Computational Intelligence.
- [5] Moustafa, N., & Slay, J. (2015). *The Evaluation of Network Anomaly Detection Systems: Statistical Analysis of the UNSW-NB15 Dataset*. IEEE