ISSN: 0974-8571 Vol.14 No. 1 June, 2022

International Journal of Computational Intelligence in Control

Improving Digital Image Forensics Analysis Using Machine Based Ensemble Learning Techniques

Hayat Ullah[1*], Muhammad Zubair Asghar[1], Muhammad Rehan[1], RizwanUllah[1], Aamir Aftab[1]

1. Institute of Computing and Information Technology, Gomal University, Dera Ismail Khan, 29050, Pakistan.

hayyat.ullah2468@gmail.com
mzubairgu@gmail.com
mrehanmarwat790@gmail.com
rizwanullah9910@gmail.com
aamiraftab877@gmail.com
*Corresponding Author: hayyat.ullah2468@gmail.com

Date of Submission: 15th April 2022 Revised: 20th May 2022 Accepted: 2nd June 2022

ABSTRACT

Due to the advancements in the internet and technology, platforms of social media have gained popularity worldwide among users. People share their opinions, feelings and views using a multiple social media platform. That's why people are using these platforms to share manipulated and forged visual content. One common technique includes training a classification model to recognize tempering or inconsistencies in the structure of images, pixel patterns, or lighting. Multiple Machine learning and deep learning algorithms, like (CNNs) Convolutional Neural networks, are often utilized for this task. However, the existing works are deficient in terms efficient detection of forged images due to inherit problems associated with classical machine learning models. To overcome these issues, we propose an efficient ensemble learning technique for efficient detection of forged images. The proposed work is aimed at developing a forgery detection system to detect manipulated images. A features extraction module is employed to extract visual feature by ensemble learning techniques respectively. Our Proposed ensemble learning approachcombines the output of various Machine learning classification models and attains a classification accuracy of 93%, significantly outperforming the individual models.

Keywords: Ensemble Learning, Machine Learning, Forgery images, Stacking technique, Not Operator.

1. INTRODUCTION

1.1 Introduction and Background

In daily life, images contain everything about people and briefly present the courtroom in every field as evidence, such as Newspapers, magazines, education, medical diagnosis and entrainments. With the advanced tools used for image editing of software like ImageJ and Photoshop, digital images can easily be tempered with any traces left, and judicial sentences,

Medical images, insurance claims and patent infringements can be affected by tempered images. As advancement image processing technology, without any clue or visual sign the digital image is tempered [1]. On social media, fake images will cause disturbance and communal tension in society [2].based on the image's method [3] focusing on frequency methods [4] relying on pattern analysis. The aim of this work is to develop a Machine learning ensemble-based system to identify and predict Forgery classification from images. Using stacking-based Ensemble learning for a proposed model where meta-model as MLP model used. For classification in Machine learning, Decision Tree, Logistics Regress and Naïve Bayes algorithm are utilized. In this study, Researchers used the Not operator for the BitMap masking technique. In addition, Researchers investigate the classification performance of Machine learning algorithms. Finally, the proposed model is compared with multiple classification models.

1.2 Research Motivation

In this current area of research [5,6], machine learning approaches are used for image forgery detection. From the digital photos, the main objective is detecting counterfeiting [6], applying a deep learning methodology; considering performance loss can be attributed to improving Machine learning ensemble techniques. However, we used the ensemble learning technique in our suggested Machine learning architectures for forgery image detection.

1.3 Problem Statement

Detection as a binary classification task, formulate the real and counterfeit image detection from forgery images as a classification problem. Given a Forgery images Class = {0 and 1} as input, the aim is to design classification models/classifiers which are used for image detection and assign corresponding Class to images. For this study, researchers applied experiments that combined multiple machines learning classifiers using an ensemble learning approach. Furthermore, the classification performance of the proposed models will be compared with different forgery image detection systems using an ensemble learning technique.

Deepfake detection also faces computation cost related challenges such as adversarial robustness, generalization and mode interpretability. Many other existing techniques rely on signal architecture model.

1.4 Research Questions

RQ1: How to apply different Machine learning-based ensemble techniques for efficient detection of forgery images?

RQ2: How to implement ensemble learning techniques for Deep fake predictions?

RQ3: What is the efficiency of the proposed ensemble model for predicting forgery images?

1.5 Research Significance and contributions:

The weak learning algorithms` classification accuracy for forgery image prediction is not efficient. A number of studies earlier in this specific area emphasize the single or hybrid use of classification algorithms (Weak learner). In this study, Researchers focus on the reasons for classifying forgery image detection from a selected dataset and conduct experiments using algorithms on the basis of ensemble techniques. Furthermore, Researchers compared the classification results of ensemble-based technique of machine learning approaches. This suggested work focuses on classifying of forgery image detection using an ensemble-based machine and deep learning technique.

The mentioned points are the most important contributions made by the suggested research:

- The suggested system employs Machine learning ensemble algorithms to perform digital Image forensics with the aim of identifying Image forgery.
- A suitable number of classifiers are contained within the Machine ensemble learning approaches.
- The performance of the proposed system compared with individual Machine Learning classifiers.
- Evaluation of the effectiveness of the proposed model in the form of comparable research.

The above-mentioned description is the layout of this article.

2. RELATED WORK

2.1. Introduction:

In this section Forgery images recognition related literature reviews to classification and detection. We have discussed and review existing approached for forgery images detection are discussed.

Several studies which used for forgery detection in Deep learning model, where CNN architecture is more effective [7, 8] found that FCN-MFCN is play a vital role while [9], highlighted CNN architecture have a ability to capture hidden features. Algorithms like Mantra-Net, Resnet, Cat-Net, Yolo-CNN and Buster-Net have been used [10, 6] utilized DenseNet-121 for JPEG compression and hit an accuracy 91.74. On CASIA and DEFACTO [11] introduced MVSS-Net, and Hit an accuracy with 88%. On the dataset of NIST developed EMT-Net for detection of Edge, and hitting 82% F1-score. In Future improvement Using GANs technique, refining detection of edge and dataset expanding for better generalization.

2.2 Research Gap

Multiple Machine learning classifiers have been applied for the detection forgery of images such as Retouching, Copy move, Watermarks, and others. However, according to existing literature, no ensemble learning technique gives such classification accuracy of 93%, which is gained by the proposed model detection of forgery images and comparison of their result between machine learning ensemble learning. Never used an MLP neural model as a Meta-learning in ensemble learning stacking technique. Never using Proposed model classification algorithms such as NB, DT, and RF as machine learning. On the CASIA dataset.

3. METHOLODOGY

3.1 Introduction.

This section emphasizes a comprehensive study of ensemble learning techniques for digital forensics, specifically in identifying forensic image instances. The primary objective is to extract forged images from visual data and classify them as real or fake. The following section provides a detailed overview of the proposed approaches.

A novel dataset was downloaded to facilitate the identification of real and fake images. Preprocessing techniques and feature extraction modules were applied to enhance visual content analysis. Additionally, multiple ensemble machine learning techniques were employed

for forgery detection. Furthermore, the performance of the proposed models was evaluated and compared with existing forgery detection systems.

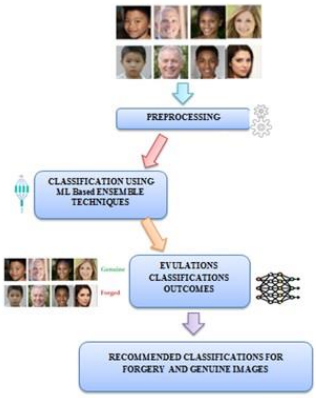

Fig1: Model of ensemble learning Technique for Forgery images

3.2 Dataset

The first step in this study is dataset acquisition to achieve the research objectives. The CASIA 2.0 benchmark dataset provides real and manipulated images for evaluating forgery detection effectiveness. It contains 12,477 images (JPG, BMP, and TIF formats), including 4,971 manipulated and 7,506 authentic images. The dataset covers various categories such as buildings, wildlife, people, trees, environments, and interiors, with resolutions ranging from 800×600 to 384×256 pixels.

Experiments were conducted on an AMD 64 system (Family 6, Model 79, Stepping 1) with a TITAN Xp Collector's Edition GPU (12GB total memory, 11.6GB free, 6% load, 35°C temperature). Data splitting follows a ternary approach, dividing it into training and test sets (see Fig. 3). Table 1 provides dataset details, and Figure 2 showcases sample images

Table 1: Depict detailed information of the CASIA 2.0 Benchmark Dataset

Training (80%)		**Testing (20%)**		**Total (images)**	
Real images	Counterfeit images	Real images	Counterfeit images	Real images	Counterfeit images
6004	3977	1502	994	7506	4971
9,981		**2,496**		**12,477**	

FIG 2. Sample snaps from CASIA dataset. 1st row (genuine images). 2nd (forged images).

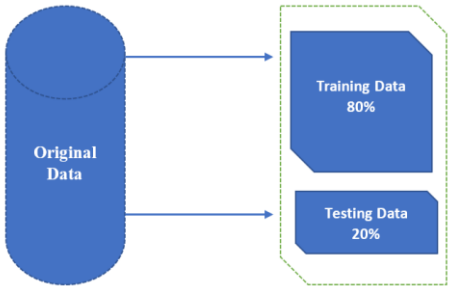

Fig 3: original data partitioning

3.2.1 Set of train data:

The training dataset is used to train the model by providing input data along with corresponding outcomes [12]. While classification models typically achieve high accuracy during training, their efficiency often declines over time. To address performance issues related to overfitting and underfitting, appropriate measures are implemented [12].

3.2.2 Set of Testing data:

A set of testing data is used to evaluate model efficiency based on a new/unobserved instance. It`s applied when the classification model is fully trained. The test part of the data performs the classification model's final estimation [12].

3.3 Pre-Processing Input Images

In this section there are multiple pre-processing techniques applied, such as converting images into Grayscale, Resizing the images, and Normalizing the pixels of images. BitMap Masking Technique using Not Operation.

3.3.1 Reading the images in Grayscales.

It contains only instance information without color, reducing the data size required for model computations. In color images, each pixel is represented by three values—Red, Green, and Blue (RGB), whereas in grayscale images, each pixel has a single intensity value. Converting RGB images to grayscale enhances computational efficiency but may impact performance accuracy by approximately 3% [13].

3.3.3 Normalized value of Pixels.

It helpsto ensure that the data input has a constant range. We normalized the pixels value and converted to floating-point numbers 255.0 in the range (1-0). It ensures that the data fed in the model is more manageable and consistent. Through normalization, neural network train fast and give more effective results.

3.4 Feature Engineering:

Selection of feature process are following

3.4.1 BitMap Masking

The CASIA dataset has masking images of Fake images in the folder name Ground Truth. It is a way which is used for computer graphic for managing the blending and visibility of different parts of images. It an uncompressed and more convenient as compared other images formats.

3.4.1.1 Masking

Masking on images are used to determined that the which images parts should be hidden or visible. It is also itself a bitmap where each and every pixels of images should be indicated that whether it should be displayed or not.

3.4.1.2 Binary masking.

Bit sampling extracts image data by representing each pixel as either 00000000 or 11111111 based on the sampled bit. For example, if a pixel is represented in binary as 10101111, 01101110, or 01001101, the least significant bit (LSB) representation would be 11111111, 00000000, or 11111111, while the most significant bit (MSB) representation could be 00000000, 00000000, or 11111111 [14].

Binary masking (0 and 1) indicates pixel visibility—1 for visible and 0 for hidden. The bitmask technique utilizes a Bitwise mask (0b11111111), where 0b signifies binary notation, and an 8-bit binary number (e.g., 255 in decimal) defines bit values.

This study employs the NOT Operation Bit Masking technique, where tampered images are processed with the NOT operator to detect forgeries using known ground truth masks. Fig. 4 illustrates the process: the first image is the original (TP folder), the second is the ground truth, and the third shows the grayscale transformation after applying BitMap masking with the NOT operation.

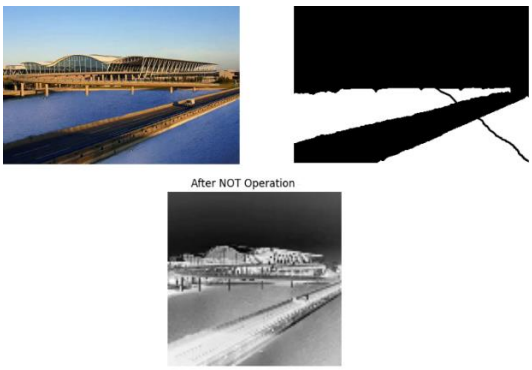

Fig 4: Display the image conversation after applying Not BitMap Operation.

3.4.2 Flatten Images / Reshaping:

Flattening images (X.reshape(len(X), -1)) simplifies data for classifiers by converting 2D arrays into 1D feature vectors [15]. 1D position embeddings, finding no advantage in 2D. This is a common preprocessing step for masking and TP class images.

3.5 Ensemble learning:

Ensemble learning combines multiple classifiers to improve decision-making, compensating for weaker models with stronger ones [16]. It has shown superior performance over single classifiers and is widely used in bioinformatics [17] and computer vision [18]. Modern applications in image classification and regression enhance accuracy [19,20]. Common ensemble methods include Averaging, Voting, Boosting, Bagging, and Stacking.

3.5.1 Stacking Classifier

It's a structured ensemble learning where used multiple layers, where first layers provide output for the 2nd layers input. Generated prediction Base classifiers which train a meta classifier and then refine a final output. Researchers used multiple classifiers as meta model as shown in Figure 5. Where Three machine learning classifiers trained individually and then their prediction train the Meta model. Stacking technique used Three different key rules such as using Sub-training data for base classification model result, Generate new prediction and then at last train the meta-classifier on these prediction result.

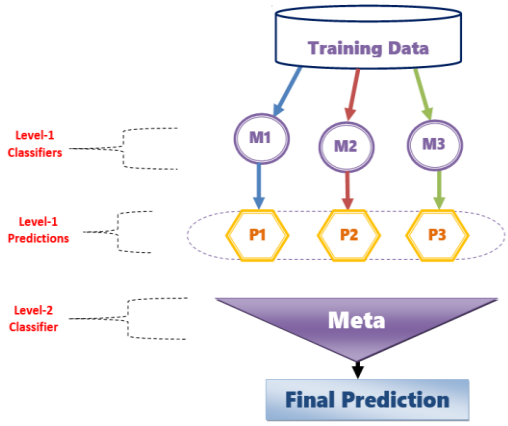

Fig 5 Stacking classifier Diagram

Stacking Classifier:

Input: Set of training data $D = \{a_i, b_i\}^k_{i=1}$ ($a_i \in X^m$, $b_i \in Y$)
Output: Ensemble Classifier S

Step-1: read level-1 classifier
For $h \leftarrow 1$ to H
do
read base-classifier s_h based on D
end for
Step-2: create other dataset from D
for
$i \leftarrow 1$ to k
do
create new dataset which holds $\{a'_i, b_i\}$, where $a'_i = \{s_1(a_i), s_2(a_i), ..., s_H(a_i)\}$
end for
Step-3: read level-2 classifier
based on the recently generated dataset read new classifier s'
return $S(a) = s'(s_1(a_i), s_2(a_i), ..., s_H(a_i))$

Algorithm 1: Show the stacking working process:

4. RESULTS AND DISCUSSION

This chapter analyzes and provides results from numerous experiments which is related to the concerned research questions.

4. 1 Address RQ 1.

Machine learning-based ensemble learning approach is able to provide more stable classification results with lower variance for a variant of causal factors as compared with individual Machine learning models. To answer Q1 Researchers applied multiple machine learning models as a Base estimator's model and MLP model as Meta model.

4.1.1 Machine Learning approaches.

Many Machine learning models apply from multiple including families such as Logistics Regression (LR), Random Forest (RF), Naïve Bayes (NB), Support Vector machine (SVM), and K-Nearest Neighbors (K-NN) with the ensemble stacking techniques. The researchers also used three Base estimator models: such as Decision tree (DT), Logistics Regression (LR), and Naive Bayes (NB).

There are several well-known meta classifiers, such as Decision tree (DT), Support Vector Machine (SVM), Random Forest (RF), Ada Boost, Light GBM, XGBoost, Cat Boost, K-Nearest Neighbors, and Multi-Layer Perception (MLP) Researchers also use a Multi-Layer perception (MLP) meta learner model.

4.1.1.1 Decision Tree:

The multiple parameter values unit of machine learning model Decision Tree is applied, as listed in Table 2.

Table 2: Having Parameters value for the Machine learning model Decision Tree.

Decision Tree	**Parameter values with their Values**
Max_depth	**10** Their set the maximum depth of model of decision tree. Limits number of nodes maximum 10 in the tree.
Random_state	**42** Here ensure that the model result is reproducible. Number is 42 where its just a random number generator
Criterion	**Gini** Used for measuring the quality of best Splitting
Splitters	**Best** Split at each node this strategy is used.
Min_sample_split	**2** Minimum number of required samples for internal node splitting.
Min_sample_leaf	**1** Ensure that leaf have one sample at least

In Table 2, the Researcher discusses the value of multiple parameters regarding Machine learning models of Decision tree. Values of maximum depth, Random state, Criterion, Splitters, Min_sample_split, and Min_sample_leaf is set as fixed.

4.1.1.2 Logistics Regression:

The multiple by defaults parameter of the machine learning model Logistics Regress listed in Table 3.

Table 3: Having the Parameters value of Machine learning model Logistics Regress.

Logistics Regression	**Parameter values with their Values**
Penalty	**12** Used for control overfitting
C	**1.0** Smaller values state stronger regularization
Solver	**1bfgs** Used for problem of optimization.
Tol	**le-4** Default value is 0.0001 Used algorithm for when change in the cost function is smaller
Multi classes	**Auto** Specific for handling multi classes
Max_iter	**100** Maximum numbers of iteration controller.
Intercept_scaling	**1** Used for controlling intercept of scaling
Warn_state	**False** Reuse the solver of last call for fitting initialization.

In the table 3, the Researcher discusses the value of multiple parameters regarding Machine learning models of Logistics Regression. Default Values of penalty, C, Solver, Tol, Multiple classes, Max_iter, Intercept_scaling, and Warn_state is set as fixed.

4.1.1.3 Naïve Bayes:

The multiple by default parameter values unit of machine learning model Naïve Bayes are listed in table 4.

Table 4: Having the Parameter value of Machine learning model Naive Bayes.

Naive Bayes	**Parameter values with their Values**
Var_smoothing	**le-9** added small variance for avoiding division by 0.
Priors	**None** Setting priors probability from training data.

In the table 4, the Researchers discusses the value of multiple parameters regarding Machine learning models of Naive Bayes. Default Priors, and Var_smotting are set as fixed.

4.1.1.4 MLP (Multilayer Perceptions):

Multiple Meta model parameters values unit of MLP (Multilayer Perceptions) are used in ensemble learning are listed in table 5.

Table 5: Having a parameters value of Meta model for using stacking ensemble learning teachings.

Multilayer Perceptions	**Parameter values with their Values**
hidden_layer_sizes	**(100,50)** It defines neurol network architecture by setting layer size 100,50, where model have 1st hidden layer 100 and 2nd one is 50 neurons.
Activation	**ReLU** Its activation function of ReLU for introduce non-linearity in model.
Solver	**Adam** weight optimizer is adam.
Alpha	**0.001** Regularization parameter also called L2 penalty used for overfitting.
Batch_size	**Auto**
Learning_rate	**Constant** learning rate remains constant
Learning_rate_init	**0.001** Used by solver initial learning rate
Power_t	**0.5**
Max_iter	**200**
Shuffle	**True** Shuffle training data before every epoch.
Early_stopping	**False** Stop training early when validation score not improve.

Validation_fraction	0.1 Floating value, training data used for validation.

In Table 5, Researchers discuss the value of multiple parameters regarding Machine learning models of Multilayer Perceptions (MLP) as a Meta learner. The Default Values of activation, hidden layer, Solver, Alpha, Batch size, learning rate, learning_rate_init, Power_t, max_iter, shuffle, early_stopping and validation_fraction are set as fixed.

4.2 Address RQ2.

In this study used ensemble learning techniques to enhance the model classification performance of Machine learning. Ensemble learning merges the decision of multiple Base estimator's model classifiers using various techniques such as averaging or voting for final decision improving [21]. It has three branches categories such as Boosting, Stacking, and Bagging. Based on meta-learning Stacking ensemble learning techniques are the best technique, where meta-learning model learn from data, how to weight Base estimators model classifiers and then combine them in the best way for optimize the classification performance of the resulting model. The Stacking technique optimizes a series of heterogeneous base classification models and marge their decision using meta learners[22].

4.2.1 Training individual Model

Every Base-estimator learner model trains on the preprocessed data and produces its classification prediction for the output at level 0 and then use the output prediction of these Base estimator model as input features for train a meta model and at the end,the meta model merges the classification prediction of Base estimator model to make the final classification report.

4.2.2 Ensemble Technique

In this study, used an ensemble stacking model learning optimization, which combines the three-machine learning Base estimators' model such as DT, LR, and NB.These diverse classifiers were Integrated v.i.a using a meta-learner (MLP) to improve the classification performance of forgery images prediction. Its role is to aggregate the classification result from the Base estimator's model and properly obtain each and each straight while mitigating their separate flaws. The Ensemble techniques are robust the efficiency of Machine learning classification algorithms for structure data handling. The main goal to get a target accuracy and reliability for the classification and detection of forgery images.

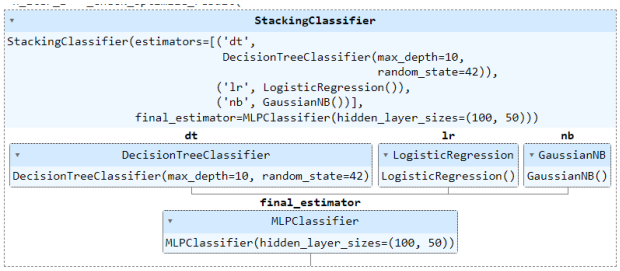

Fig 6: Show the flowchart of Ensemble learning Technique for Machine Learning

4.3 Address RQ 3.

To find an answer to RQ 3, This experiment is applied on the dataset compressing 12485 forgery images about fake and real images. Multiple machine learning algorithms, such as Decision tree, Logistics Regresion, and Naive Bayes are used as theBase estimator model due to the stacking ensemble learning technique, and the Metal model is Multi-layer Perception (MLP) model. Researchers used different metrics, including Precision, Recall, F1 score, and Accuracy. It verifies that the performance of the ensemble learning technique is much better in the way of getting evaluation metrics results with Precision 93%, Recall 93%, F1-Score 93%, and Accuracy 93%.

Also, Researchers discuss each machine learning classification algorithm`s results individual. In the Individual report, the Decision Tree Machine learning classifier plays a novel role as compared with other suggested Machine learning classifiers. Decision Tree gives precision 88%, Recall 88%, F1-Score 88%, and accuracy 88%. This is Show in table 6.

Table 6: Shows that the Individual accuracy of **Decision Tree** is 88% of the Machine learning classification model.

Classes	**Precision**	**Recall**	**F1 score**
0	91%	89%	89%
1	84%	86%	85%
Weighted avg	88%	88%	88%

Table 7 Present the accuracy result of Machine learning Classifier **Logistics Regression** which generate Accuracy 87% result are following.

Table 7: Show the classification result of Machine learning classifier Logistics Regression.

Classes	**Precision**	**Recall**	**F1 score**
0	91%	87%	89%
1	92%	87%	84%
Weighted avg	87%	87%	87%

Table 8: Present the Classification Report of Machine learning algorithm **Naive Bayes** where its accuracy is 84% and Precision, Recall, and F1- Score are following.

Classes	**Precision**	**Recall**	**F1 score**
0	89%	84%	86%
1	78%	84%	81%
Weighted avg	84%	84%	84%

In the below table researchers show the classification result of Ensemble based Machine learning algorithm using stacking technique.

Table 9 Depict the Experimental results of Machine based Ensemble learning technique.

Classes	**Precision**	**Recall**	**F1 score**
0	93%	94%	96%
1	93%	89%	91%
Weighted avg	93%	93%	93%

O class is used for Real images, and 1 class is used for fake images because there are 2 types of images classes, such as Real and Fake. The Confusion matrix table follow. There are two folders in the dataset namely (Au, and Tp). Au is used for Real images and Tp is used for fake images.

4.3.1 Comparisons of individual Machine learning model with Ensemble Machine learning Technique:To estimate the proposed ensemble model technique performance with individual machine learning model techniques (NB, LR, DT) for Deep fake prediction, Researchers perform experimentation. The Naïve Bayes model in Table 10 gives a Poor classification report in regarding with multiple evaluation metrics such as (Recall 84%, Precision 84%. F1-Score 84% and Accuracy 84%). The Decision Tree classification model in Table 10 gives a performance classification report in regarding with multiple evaluation metrics such as (Recall 88%, Precision 88%, F1-Score 88%, and Accuracy 88%). The classification model Logistics Regression in table 10 depicts a classification report regarding with multiple evaluation metrics such as (Recall 87%, Precision 87%, F1-Score 87%, and Accuracy 87%).

The above discussed results show the result of Machine learning ensemble learning techniques when a comparison is performed with Multiple Machine learning models such as individual Decision Tree, Logistics Regression, and Naïve Bayes. The ensemble learning

techniques model better performance classification result in table 10, i.e (Recall 93%, Precision 93%, F1-Score 93%, and Accuracy 93%).

Table10: Depict the comparison of Ensemble Technique Machine learning performance with Individual machine learning models result

Method	Accuracy (%)	Precision (%)	Recall (%)	F-score (%)
DT	88	88	88	88
LR	87	87	87	87
NB	84	84	84	84
Machine learning Ensemble learning	**93**	**93**	**93**	**93**

5. CONCLUSION AND FUTURE WORK

A system capable of analyzing has be proposed for data to predict and identify. Our proposed system consists of different modules for forgery images categorized into real and fake classes. Multiple modules have our proposed system like 1) Forgery images detection development, 2) Data partitioning, 3) Pre-Processing data, 4) Bitwise mapping Not Operation, 5) Features extraction, 6) Deep fake images detection, 7) Ensemble learning technique using machine learning model.

We have acquired a forgery images benchmark dataset (CASIA v2) from the Kaggle site. The novel dataset, consisting of 12,477 total images, 7,506 of which are real images and 4,971 are counterfeit, is used for the proposed work.

The forgery images were splitting into testing and training sets. Both types of sets were further used in the pre-processing module. In this pre-processing module, images content was pre-processed using multiple strategies, such as image resizing, grayscale conversion, and a Bitmap masking technique. Additionally, features were extracted for pre-processed images. Multiple architectures were used to acquire visual image. Furthermore, we have optimized the proposed model performance (Ensemble learning techniques using Machine learning models).

We have conducted multiple investigation for forgery images detection. The results show that the proposed model of ensemble learning techniques using deep learning models outperformed as comparing other models with a classification accuracy of 93%.

5.1 Limitation:

The following limitations could be considered for this research.

- Forgery analysis can improve using DT+LR and SVM+NB.
- Only stacking based ensemble learning technique is used, ensemble learning others techniques are not used.
- Convert images into grayscale formats.
- Bitmap masking Not operation is utilized another Bitmap operation are not used.

5.2 Future work:

- Combine images with Audio, Video for comprehensive analysis.
- Proposed model modification in specific domain like, E-commerce, Healthcare and Agriculture etc.
- RGB images can also affect on the classification of proposed model.

- Bitmap Masking operation such as AND, OR, NOR operation can also affect on proposed model classification result.
- Other stacking-based learning technique can be utilized for proposed work.

Conflict of Interest

Authors declare that they have no competing interests.

Authors' Contributions

All authors contributed equally.

Authors Contributions

- H Ullah, and RUllah performed data curation, H Ullah, M Rehan conceived experiments, HUllah, RUllah, A Aftab compiled results and reviewed the paper, M Z Asghar performed overall supervision and wrote the paper.

REFERENCES

[1]. Popescu, A. C., & Farid, H. (2004, May). Statistical tools for digital forensics. In International workshop on information hiding (pp. 128-147). Berlin, Heidelberg: Springer Berlin Heidelberg.
[2]. Mahmood, T., Mehmood, Z., Shah, M., & Saba, T. (2018). A robust technique for copy-move forgery detection and localization in digital images via stationary wavelet and discrete cosine transform. Journal of Visual Communication and Image Representation, 53, 202-214.
[3]. Chai, L., Bau, D., Lim, S. N., & Isola, P. (2020). What makes fake images detectable? understanding properties that generalize. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVI 16 (pp. 103-120). Springer International Publishing.
[4]. Frank, J., Eisenhofer, T., Schönherr, L., Fischer, A., Kolossa, D., & Holz, T. (2020, November). Leveraging frequency analysis for deep fake image recognition. In International conference on machine learning (pp. 3247-3258). PMLR.
[5]. Kuznetsov, A. (2019, November). Digital image forgery detection using deep learning approach. In Journal of Physics: Conference Series (Vol. 1368, No. 3, p. 032028). IOP Publishing.
[6]. Ali, S. S., Ganapathi, I. I., Vu, N. S., Ali, S. D., Saxena, N., & Werghi, N. (2022). Image forgery detection using deep learning by recompressing images. Electronics, 11(3), 403.
[7]. Salloum, R., Ren, Y., & Kuo, C. C. J. (2018). Image splicing localization using a multi-task fully convolutional network (MFCN). Journal of Visual Communication and Image Representation, 51, 201-209.
[8]. Barad, Z. J., & Goswami, M. M. (2020, March). Image forgery detection using deep learning: a survey. In 2020 6th international conference on advanced computing and communication systems (ICACCS) (pp. 571-576). IEEE.
[9]. Agarwal, R., Khudaniya, D., Gupta, A., & Grover, K. (2020, May). Image forgery detection and deep learning techniques: A review. In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1096-1100). IEEE.
[10]. Qazi, E. U. H., Zia, T., & Almorjan, A. (2022). Deep learning-based digital image forgery detection system. Applied Sciences, 12(6), 2851.
[11]. Chen, X., Dong, C., Ji, J., Cao, J., & Li, X. (2021). Image manipulation detection by multi-view multi-scale supervision. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 14185-14193).
[12]. Asghar, M. Z., Subhan, F., Imran, M., Kundi, F. M., Shamshirband, S., Mosavi, A., ... & Varkonyi-Koczy, A. R. (2019). Performance Evaluation of Supervised Ma chine Learning Techniques for Efficient Detection of Emotions from Online Con tent. arXiv preprint arXiv:1908.01587.
[13]. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of big data, 6(1), 1-48.
[14]. Grantham, B. (2007). Bitmap Steganography: An Introduction. COT 4810: Topics in Computer Science Dr. Dutton.

[15]. Wang, Y., Hassebrook, L. G., & Lau, D. L. (2010). Data acquisition and processing of 3-D fingerprints. IEEE Transactions on Information Forensics and Security, 5(4), 750-760.
[16]. Tang, E. K., Suganthan, P. N., & Yao, X. (2006). An analysis of diversity measures. Machine learning, 65, 247-271.
[17]. Yang, P., Hwa Yang, Y., B Zhou, B., & Y Zomaya, A. (2010). A review of ensemble methods in bioinformatics. Current Bioinformatics, 5(4), 296-308.
[18]. Li, X., Yang, H., He, J., Jha, A., Fogo, A. B., Wheless, L. E., ... & Huo, Y. (2021, April). Beds: Bagging ensemble deep segmentation for nucleus segmentation with testing stage stain augmentation. In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (pp. 659-662). IEEE.
[19]. Zheng, H., Zhang, Y., Yang, L., Liang, P., Zhao, Z., Wang, C., & Chen, D. Z. (2019, July). A new ensemble learning framework for 3D biomedical image segmentation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 5909-5916).
[20]. Yang, Y., Hu, Y., Zhang, X., & Wang, S. (2021). Two-stage selective ensemble of CNN via deep tree training for medical image classification. IEEE Transactions on Cybernetics, 52(9), 9194-9207.
[21]. Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. Wiley interdisciplinary reviews: data mining and knowledge discovery, 8(4), e1249.
[22]. Rajagopal, S., Kundapur, P. P., & Hareesha, K. S. (2020). A stacking ensemble for network intrusion detection using heterogeneous datasets. Security and Communication Networks, 2020(1), 4586875.