Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

Design of WhatsApp Image Folder Categorization Using CNN Method in the Android Domain

R. Asokan[1], T. Vijayakumar[2]

[1]Professor and Principal, Kongunadu College of Engineering and Technology, Tholurpatti, Tamil Nadu, India
[2]Professor, Department of CSE, Jai Shriram Engineering College, Avinashipalayam, Tamil Nadu, India

E-mail: [1]asokece@yahoo.com, [2]vishal_16278@yahoo.co.in

Abstract

Recently, the use of different social media platforms such as Twitter, Facebook, and WhatsApp have increased significantly. A vast number of static images and motion frame pictures posted on such platforms get stored in the device folder making it critical to identify the social network of the downloaded images in the android domain. This is a multimedia forensic job with major cyber security consequences and is said to be accomplished using unique traces contained in picture material (SNs). Therefore, this proposal has been endeavoured to construct a new framework called FusionNet to combine two well-established single shared Convolutional Neural Networks (CNN) to accelerate the search. Moreover, the FusionNet has been found to improve classification accuracy. Image searching is one of the challenging issues in the android domain besides being a time-consuming process. The goal of the proposed network's architecture and training is to enhance the forensic information included in the digital pictures shared on social media. Furthermore, several network designs for the categorization of WhatsApp pictures have been compared and this suggested method has shown better performance in the comparison. The proposed framework's overall performance was measured using the performance metrics.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021
Copyright © 2021 Inventive Research Organization

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

Keywords: deep learning, image classification

1. Introduction

The development of social media platforms has shown that people consume information at a much faster rate and in an increased volume. The lines between the creation and dissemination of information are rapidly dissolving. Even though social networks have made excellent information more widely available, their highly decentralized and uncontrolled environment allows the rapid spread of false information in large numbers [1-4]. To influence public opinion, damage people's popularity, organizations, or social groupings economically or politically, misinformation is generally produced with malevolence. Furthermore, false information travels on social media at a quicker rate, goes deeper, and reaches a wider audience. Humans are also unable to tell the difference between fact and fiction owing to the overwhelming amount of information we are exposed to on social media [5-8].

CG images are a very high pixel-contained visual quality that is developed through the advancement of image processing methods for modeling purposes. A high degree of photorealism makes it impossible to tell whether the pictures were taken with digital cameras or generated with computer graphics rendering. Even though rendering methods make our lives easier, they may also put our forensic systems in jeopardy. Figure 1 depicts a few examples of potential picture categories.

Thus, there is a significant research issue in picture forensics: differentiating between computer-generated and natural imagery. This forensic difficulty with computer-generated images may be referred to as a CG problem [9].

Many researchers have suggested different approaches to this forensic issue over the last few decades, including those that use hand-crafted features and CNNs. The dimensional features

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

are extracted to formulate the optimization problem in digital image processing for estimation mapping. Because CNN has such a high learning capacity, CNN-based techniques are used frequently and tends to be more effective in forensic investigations. On the other hand, the existing techniques ignored the blind detection issue (also known as the generalization problem) if we train a CNN model on CG images produced by "known" computer graphics rendering methods and then test it on images generated by "unknown" rendering techniques [10-13].

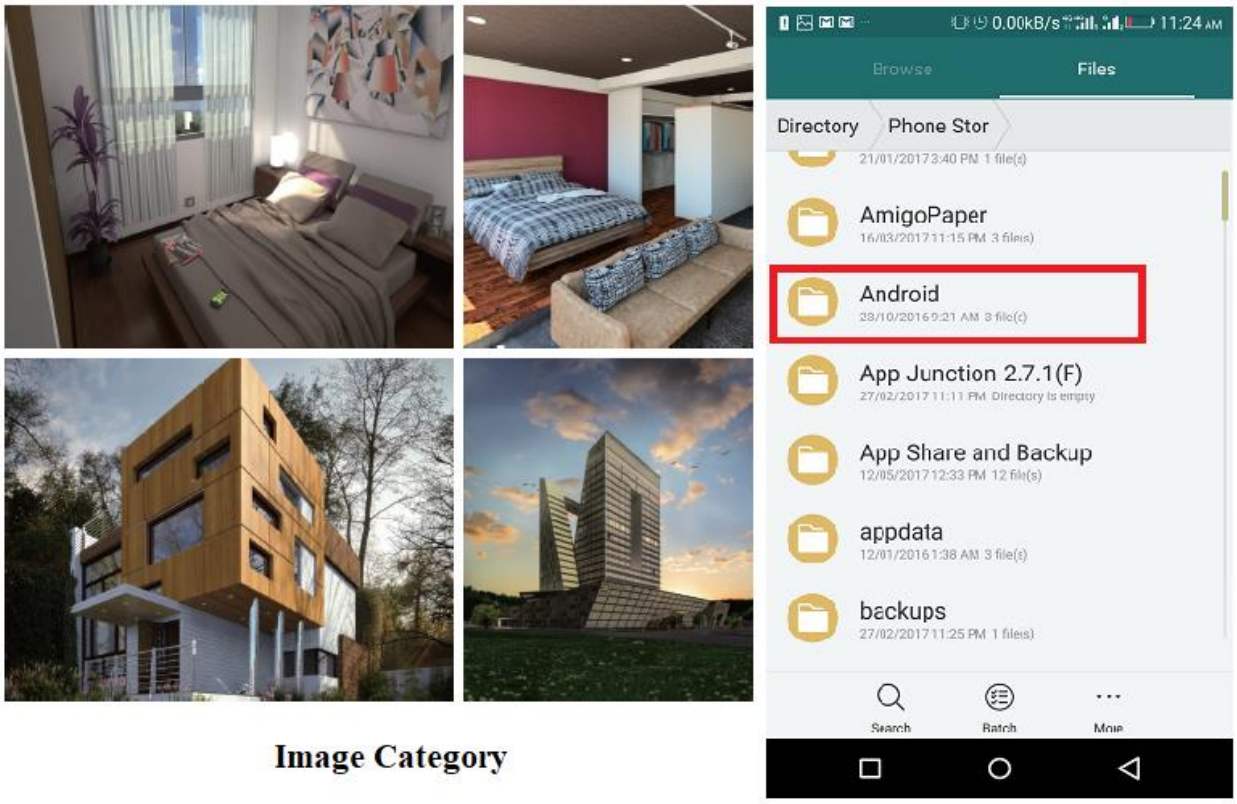

Figure 1. Sample image category

Inferring the source of digital information is a major challenge in multimedia forensics. Linking a digital picture to its originating device is unquestionably a difficult issue for researchers,

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

as one has to retrieve the full image processing history when an image has to be verified. These tasks have been made easier with the development of applications including android domain's folders to recover the various inference devices' sources [14].

2. Organization of this Research

The structure of the remaining sections of the research paper: section 3 presents previous research on image classification using a deep learning method, and section 4 presents future research on image classification. Section 4 covers the suggested technique for WhatsApp picture categorization using the CNN methodology, which describes preliminary works in detail in Section 3. Section 5 presents the experimental findings as well as the discussion of the conclusion reached in the experiment. Section 6 brings the planned work to a close, with suggestions for further improvement.

3. Preliminaries

CG pictures were identified using a CNN-based framework described by Quan et al, which included two cascaded convolutional layers at the beginning of the framework. Several CNN visualization tools were used to comprehend the differences between these two types of images by the deep model [15]. He and his colleagues used a combination of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to identify CG pictures [16]. For the CG picture forensic issue, Nguyen et al. used the capsule network technique [17]. Using an attention-based recurrent model, Bhalang Tarianga et al. developed a method to classify computer-generated pictures and natural pictures [18]. With the use of channel and pixel correlation information, Zhang et al. were able to develop a hybrid-correlation-based CNN model that exposed the differences in features between NIs and computer-generated pictures [19].

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

Using a K-NN classifier, the researchers in [20] were able to distinguish between social media to trace the image changes, feature capturing, and operations with uploading and downloading.

Caldelli and his research team have proposed classification methods for different social media networks based on discrete cosine transform (DCT). This procedure was performed with content-based feature extraction in the given input images by the decision tree classification method [21].

The combination of dimension and quantization of image details have provided a good structure for the procedure to differentiate between the features of the image which has content-based features. According to [22], in order to identify the objects, they utilized three different classifiers, namely: the LR, the SVM, and the RF classifiers.

A CNN deep learning method presented by Amerini et al., proposes to categorize up to several social media networks based on the features of the images through transform coefficients [23].

Research Gap

This hypothesis recommends to build the framework using two elements of CNN: network design and network training, to enhance forensic performance in particular. With this approach, the goal is to create and deploy a CNN that uses more diverse feature learning and uses tougher negative examples in the so-called “enhanced training phase”. As a result, the negative sample refers to the picture that was generated entirely from the original training dataset (perhaps supplemented by information from the CNN model), and its ground-truth label is identical to the CG image.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

4. Methodologies

The proposed method consists of the following sequence of layer filters for CNN construction with FusionNet.

1) This is followed by 32 3*3 filters in two convolutional layers, the last of which is then dropped out.

2) After max-pooling, dropout, and another round of 64 3*3 filters, conv3 and conv4 are constructed.

3) There are two fully linked layers (full5 and full6), each with 256 units, and a softmax layer that has as many units as class k units to be recognized.

4.1 Multiclass Importance

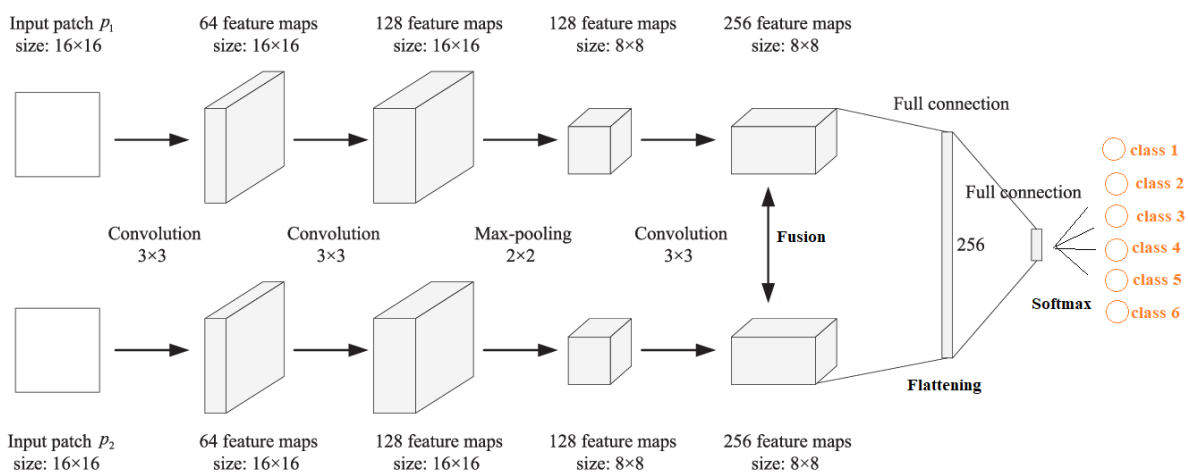

Figure 2. Proposed architecture

When dealing with a multi-class issue, we'll get a probability that a picture belongs to a certain class, as in a single dimension of CNN. The ReLU is used as an activation function in all

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

convolutional layers as well as the first fully connected layer with class 1 to 6. Among the categories of photos are screenshots, pictures of the owner (or authorized person), pictures of the family, pictures taken outside and inside, and pictures of WhatsApp memes. Increasing the number of classes in a multiclass classification system will be possible with this option. The proposed architecture is shown in Figure 2.

4.2 Training of CNN Network

The training of the CNN network has a gradient descent approach for optimizing the algorithm to improve patch or image-level training in the CNN [25]. The maximum number of epochs in a single CNN dimension is 20, whereas the maximum number of epochs in a CNN net with two dimensions is 50. For the single-dimensional CNN, a mini-batch of 32 samples was used, whereas, the two-dimensional CNN was attempted with a patch and image level of 256 pixels. An imbalanced training set was used to generate each mini-batch. Depending on the picture size, the number of photos in each class varied to better mimic real-world situations. The loss function on the testing set has attained minimum due to the training with FusionNet architecture. It was found that for the two-dimensional CNN net, loss function occured after many epochs, while for the FusionNet and the single-dimensional CNN, loss function occured before 10 epochs.

4.3 FusionNet Architecture

The static images were divided into various dimension matrixes for the patching process of the proposed architecture. As a result of this approach, after processing each picture patch using CNNs, the prediction was produced at the patch level [24]. Therefore, a majority voting method on the labels given to each patch was used to generate a final identification at the picture level. To build the suggested framework, FusionNet, the two CNNs shown in Figure 2 were combined. At any level before softmax, the fusion may take place in order to conduct a combined learning

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

process to get information from neurons in the network structure. Moreover, their fusion occurred to complete the concatenating process in the proposed architecture.

4.4 Activation Function

The neurons are combined with the activation function to drop out for the further process in the neural network structure. A second fully connected layer and a softmax with the same number of units linked to it as the number of classes to be classified follow. By using this architecture, both branches of the neural network may train at the same time using the same picture patch as input. Both branches of FusionNet share the weights throughout the training phase to automatically discover the optimum combination [25-27].

5. Results Discussion

In this section, the experimental procedure and testing and training procedure is discussed with the proposed algorithm. The experimental test is combined with feature learning and neural network architecture. The model loss and accuracy has been shown in figure 3 for training and testing.

These four datasets have already been used in relevant research when the trials were carried out. In terms of picture kinds, social network numbers and types, number of source cameras engaged, and the number of shares, all these datasets show quite distinct features [28- 32]. Table 1 displays the parameters that was used to build up this suggested model.

The VISION collection yielded a total of 2135 pictures, of which 5 were linked to specific cellphones, such as the Samsung Galaxy S3 and Huawei models. As a consequence, a total of 6405 (2135_3) pictures was downloaded in various quality levels from Facebook and WhatsApp. Since

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

the camera resolution of the mobile phone affects the size of the image, the number of samples per device varies.

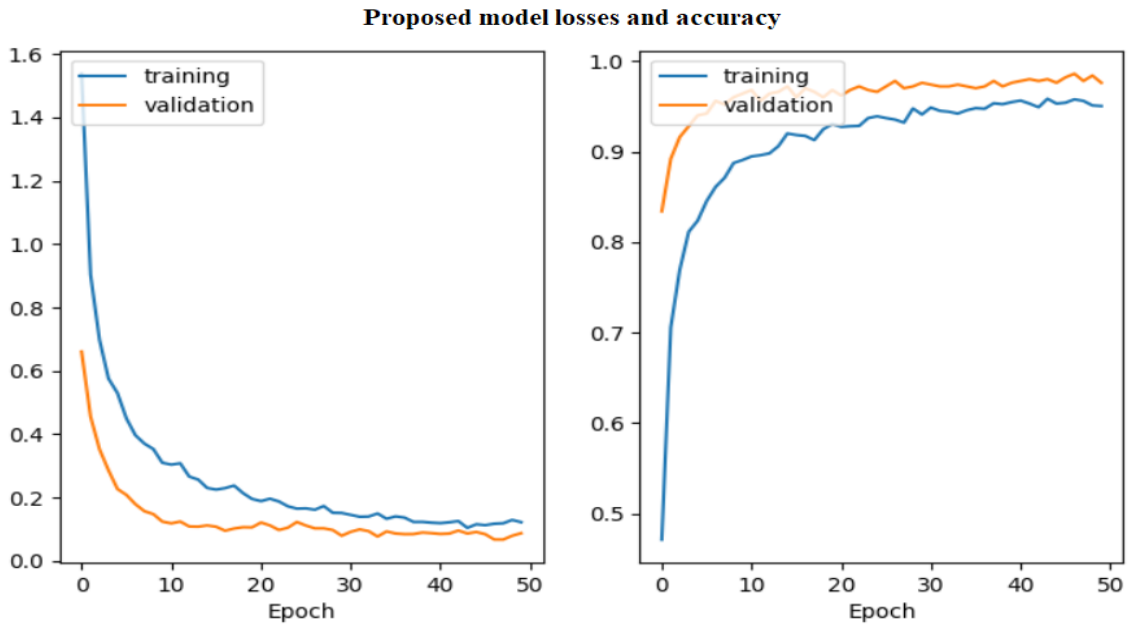

Figure 3. Training and Testing Accuracy Vs Losses

Table 1. Initial Parameters for Proposed Model

S.No	Parameter	Fusion Net
1	Initial Learning rate	0.004
2	Activation function	ReLU
3	Batch size	10
4	Epochs	50
5	Loss function	RMSE (Root Mean Square Error)

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

Table 2. Computed Performance Metrics

S.No	CNN Architectures	Accuracy	Sensitivity	Precision	Loss	Specificity
1	CNN with FusionNet (Patchlevel)	83.91%	81.67%	78.98%	2.2233	80.92%
2	CNN with FusionNet (Image level)	84.34%	80.12%	83.23%	1.8761	89.12%
3	CNN with FusionNet (DS)	74.56%	80.92%	79.06%	4.2311	81.67%
4	Proposed CNN with FusionNet (SS)	98.74%	97.47%	95.03%	0.0345	97.76%

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Specificity = \frac{TN}{TN + FP}$$

Table 2 includes performance metrics that have been computed. To demonstrate the effectiveness of the proposed approach on various datasets made up of various picture resolutions, social networks, and messaging apps, an extensive collection of experiments are given in this part. 80% of the datasets have been used for training, 10% have been used for validation, and 10% have been used for testing. Random images were chosen from each of the six groups. Graphical depiction of total performance metrics is shown in Figure 4.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

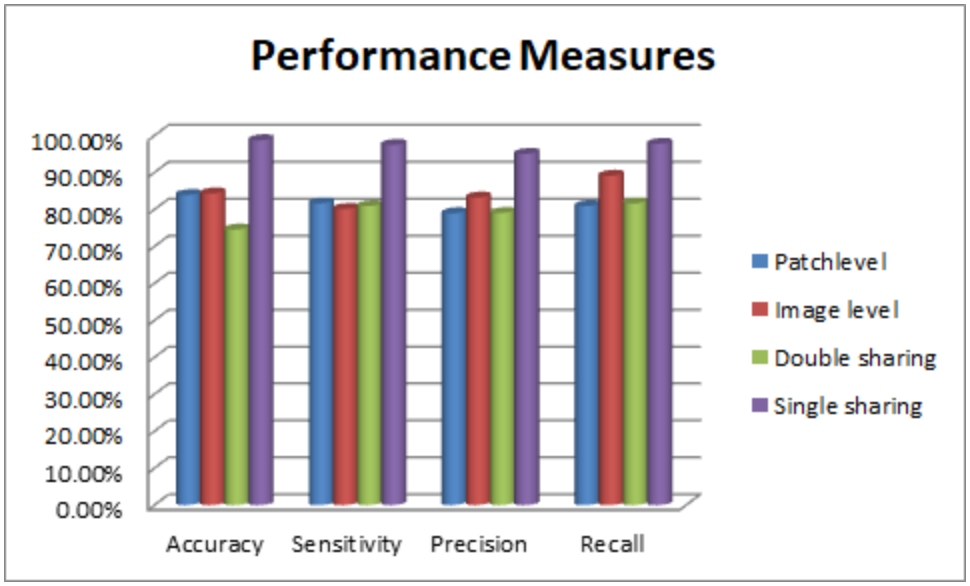

Figure 4. Overall Performance Measures

The suggested single-shared FusionNet performs more accurately than previous CNN methods, which is a significant improvement.

6. Conclusion

As a result, the proposed single sharing CNN method is very appropriate for categorizing images for simple search, and it has been successfully implemented. Using WhatsApp to disseminate false information is a significant societal issue. The accuracy of Fusion Net was superior to other conventional picture categorization techniques, as shown in the results and discussion section. In the future, this study intends to focus on combining message information (such as sender, timestamp, and group) with textual characteristics to better classify messages. A Deep Leaning method would be used to extract features from both text and media data in order to study multi-modal misinformation detection. Subsequently, the research urges to look at semi-

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

automatic techniques for creating constantly tagged WhatsApp datasets because misinformation changes over time.

References

[1] Karuppusamy, P. "Building Detection using Two-Layered Novel Convolutional Neural Networks." Journal of Soft Computing Paradigm (JSCP) 3, no. 01 (2021): 29-37.

[2] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," Journal of Machine Learning Research, vol. 11, no. Dec, pp. 3371–3408, 2010.

[3] Manoharan, J. Samuel. "Capsule Network Algorithm for Performance Optimization of Text Classification." Journal of Soft Computing Paradigm (JSCP) 3, no. 01 (2021): 1-9.

[4] G. Schaefer and M. Stich, "UCID - an uncompressed colour image database," in Proceedings of the Storage and Retrieval Methods and Applications for Multimedia, 2004, pp. 472–480.

[5] Manoharan, Samuel, and Narain Ponraj. "Analysis of Complex Non-Linear Environment Exploration in Speech Recognition by Hybrid Learning Technique." Journal of Innovative Image Processing (JIIP) 2, no. 04 (2020): 202-209.

[6] J. Yang, G. Zhu, and Y.-Q. Shi, "Analyzing the effect of jpeg compression on local variance of image intensity," Trans. Img. Proc., vol. 25, no. 6, pp. 2647–2656, Jun. 2016. [Online]. Available: http://dx.doi.org/10.1109/TIP.2016.2553521.

[7] Hamdan, Yasir Babiker. "Faultless Decision Making for False Information in Online: A Systematic Approach." Journal of Soft Computing Paradigm (JSCP) 2, no. 04 (2020): 226-235

[8] L. Van Der Maaten, "Accelerating t-sne using tree-based algorithms," J. Mach. Learn. Res., vol. 15, no. 1, pp. 3221–3245, Jan. 2014.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

[9] Dhaya, R. "Hybrid Machine Learning Approach to Detect the Changes in SAR Images for Salvation of Spectral Constriction Problem." Journal of Innovative Image Processing (JIIP) 3, no. 02 (2021): 118-130.

[10] B. Tondi, "Pixel-domain adversarial examples against cnn-based manipulation detectors," Electronics Letters, vol. 54, pp. 1220–1222(2), October 2018. [Online]. Available: http://digital-library.theiet.org/content/journals/10.1049/el.2018.6469

[11] Sungheetha, Akey, and Rajesh Sharma. "Transcapsule model for sentiment classification." Journal of Artificial Intelligence 2, no. 03 (2020): 163-169.

[12] D. Shullani, M. Fontani, M. Iuliani, O. A. Shaya, and A. Piva, "VISION: a video and image dataset for source identification," EURASIP Journal on Information Security, vol. 2017, no. 1, p. 15, Oct 2017. [Online], Available: https://doi.org/10.1186/s13635-017-0067-2

[13] Sungheetha, Akey, and Rajesh Sharma. "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network." Journal of Trends in Computer Science and Smart technology (TCSST) 3, no. 02 (2021): 81-94.

[14] F. de O. Costa, M. A. Oikawa, Z. Dias, S. Goldenstein, and A. R. de Rocha, "Image phylogeny forests reconstruction," IEEE Transactions on Information Forensics and Security, vol. 9, no. 10, pp. 1533–1546, Oct 2014.

[15] W. Quan, K. Wang, D.-M. Yan, X. Zhang, Distinguishing between natural and computer-generated images using convolutional neural networks, IEEE Transactions on Information Forensics and Security 13 (11) (2018) 2772-2787.

[16] P. He, X. Jiang, T. Sun, H. Li, Computer graphics identification combining convolutional and recurrent neural networks, IEEE Signal Processing Letters 25 (9) (2018) 1369-1373.

[17] H. H. Nguyen, J. Yamagishi, I. Echizen, Capsule-forensics: Using capsule networks to detect forged images and videos, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2019, pp. 2307-2311.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

[18] D. Bhalang Tarianga, P. Senguptab, A. Roy, R. Subhra Chakraborty, R. Naskar, Classi_cation of computer generated and natural images based on efficient deep convolutional recurrent attention model, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 146-152.

[19] R. Zhang, W. Quan, L. Fan, L. Hu, D.-M. Yan, Distinguishing computer generated images from natural images using channel and pixel correlation, Journal of Computer Science and Technology 35 (3) (2020) 592-602.

[20] M. Moltisanti, A. Paratore, S. Battiato, and L. Saravo, “Image manipulation on facebook for forensics evidence,” in Image Analysis and Processing — ICIAP 2015, V. Murino and E. Puppo, Eds. Cham: Springer International Publishing, 2015, pp. 506–517.

[21] R. Caldelli, R. Becarelli, and I. Amerini, “Image origin classification based on social network provenance,” IEEE Transactions on Information Forensics and Security, vol. 12, no. 6, pp. 1299–1308, June 2017.

[22] Q. Phan, C. Pasquini, B. G., and D. N. F., “Identifying image provenance: An analysis of mobile instant messaging apps,” in Proc. MMSP 2018, 2018.

[23] I. Amerini, T. Uricchio, and R. Caldelli, “Tracing images back to their social network of origin: A cnn-based approach,” in 2017 IEEE Workshop on Information Forensics and Security (WIFS), Dec 2017, pp. 1–6.

[24] Tripathi, Milan. "Analysis of Convolutional Neural Network based Image Classification Techniques." Journal of Innovative Image Processing (JIIP) 3, no. 02 (2021): 100-117.

[25] A. Bharati, D. Moreira, A. Pinto, J. Brogan, K. Bowyer, P. Flynn, W. Scheirer, and A. Rocha, “U-phylogeny: Undirected provenance graph construction in the wild,” in 2017 IEEE International Conference on Image Processing (ICIP), Sept 2017, pp. 1517–1521.

[26] Sungheetha, Akey, and Rajesh Sharma. "A Novel CapsNet based Image Reconstruction and Regression Analysis." Journal of Innovative Image Processing (JIIP) 2, no. 03 (2020): 156-164.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

[27] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, “Determining image origin and integrity using sensor noise,” IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, pp. 74–90, 2008.

[28] Kaur, Jasmeet, and Anil Kumar. "Speech Emotion Recognition Using CNN, k-NN, MLP and Random Forest." In Computer Networks and Inventive Communication Technologies, pp. 499-509. Springer, Singapore, 2021.

[29] Jaiswal, Akshat, Prashanth Duvvada, and Lekha S. Nair. "Hybridization of TrellisNet with CNN." In Computer Networks and Inventive Communication Technologies, pp. 769-778. Springer, Singapore, 2021.

[30] Bhandi, Vijayakumar, and KA Sumithra Devi. "Feature Extraction from Ensemble of Deep CNN Model for Image Retrieval Application." In Data Intelligence and Cognitive Informatics, pp. 725-738. Springer, Singapore, 2021.

[31] Chincholkar, Shruti, and Manoov Rajapandy. "Fog image classification and visibility detection using CNN." In International Conference on Intelligent Computing, Information and Control Systems, pp. 249-257. Springer, Cham, 2019.

[32] Hu, Fanghong, Zhu Li, and Liping Yan. "CNN and Raspberry PI for Fruit Tree Disease Detection." In International Conference on Intelligent Computing, Information and Control Systems, pp. 1-8. Springer, Cham, 2019.

Author's biography

R. Asokan received his B.E degree in electronics and communication from Bharathiar University and MS degree in electronics and control from Birla Institute of Technology. He obtained an M.Tech degree in electronics and communication from Pondicherry Engineering College, with distinction. He obtained PhD in information and communication engineering from Anna University, Chennai. He is currently the Principal, Kongunadu College of Engineering and

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2021)
Vol.03/ No.03
Pages: 180-195
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2021.3.003

Technology, Thottiyam, TamilNadu, India. He has published more than 65 papers in national and international journals and conferences. He has over 25 years of teaching experience. He is a member of various scientific and professional societies. His areas of interest include wireless networks, network security and image processing.

T. Vijayakumar is currently working as Professor in the Department of ECE at Jai Shriram Engineering College, Avinashipalayam, Tamil Nadu, India. His research includes Computer Vision, Motion Analysis, Stereo Vision, Object Recognition, computer graphics, photo interpretation, image retrieval, Embedded Image Processing and Real-time image and video processing applications.

ISSN: 2582-337X
Submitted: 02.08.2021
Revised: 14.09.2021
Accepted: 06.10.2021
Published: 16.10.2021