International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Yakshagana Character Identification Through Deep Learning with Crown and Facial Makeup Pattern Analysis

Anantha Murthy [1], Sanjeev Kulkarni [2], Amrita Arjun Kindalkar[3], Mamatha S J[4]

Submitted: 04/02/2024 **Revised**: 12/03/2024 **Accepted**: 18/03/2024

Abstract: Yakshagana, an intricate theatrical art form originating from Karnataka, encompasses variations such as Thenku-thittu, BadaguThittu, and Badaabadagu Thittu. This research delves into the historical roots, contemporary influence, and evolving makeup trends within Yakshagana. Within the Tenkutittu Yakshagana, diverse crown types take center stage, with the performer's chosen crown and facial makeup pattern serving as key determinants of the portrayed character. Our study focuses on character classes including Vishnu, Devi, Sarpa and Mahisha for character identification. To address the intricate task of character classification in Yakshagana images, this paper employs deep learning methods such as Three Tier CNN and YOLOv5 . Specifically, a Cyclic Gate Recurrent Neural Network is utilized to classify characters like Vishnu, Devi, Sarpa and Mahisha. Following character categorization, the model proceeds to determine disguises. The Three-tier CNN achieves a commendable 90% accuracy in classifying disguises. Through thorough testing, it has been established that YOLOv5, boasting a remarkable 95% accuracy in identifying multiple elements within an image, emerges as the most suitable algorithm for character identification. This research serves as a real-time tool, aiding newcomers in identifying the appropriate crown and makeup pattern for specific Yakshagana figures.

Keywords: *character identification, Three tier Convolution Neural Network (CNN), YOLOv5 , Cyclic Gate Recurrent Neural Network (CGRN), Object detection, Yakshagana*

1. Introduction

Yakshagana finds its roots in ancient folk traditions, with influences from various sources including mythological narratives, religious rituals, and cultural practices. Historically, it is believed to have originated in the 16th century during the Bhakti movement, gaining prominence as a form of devotional theatre. The term "Yakshagana" itself is derived from "Yaksha" (nature spirits) and "Gana" (song), indicating its association with folk tales and mythological characters. Yakshagana, a traditional form of theatre originating from Karnataka, India, faces many challenges in the modern period. Despite its rich cultural heritage and historical significance, Yakshagana grapples with various issues that threaten its continuity and relevance. It explores the current challenges confronting Yakshagana, shedding light on the complexities of preserving tradition in a rapidly changing world. One of the foremost challenges facing Yakshagana is the economic strain on its practitioners and organizers [2]. The rising costs of production, including costumes, makeup, and musical instruments, pose financial barriers for traditional Yakshagana troupes. Changing audience demographics present another significant challenge for Yakshagana. With the younger generation showing a preference for digital entertainment and Westernized cultural influences, there is a decline in interest and participation in traditional art forms like Yakshagana.. The absence of robust institutional support poses a critical challenge to the sustainability of Yakshagana [12]. Unlike other mainstream art forms, Yakshagana lacks formal recognition and support from government agencies and cultural institutions. This lack of institutional backing hampers efforts to preserve and promote Yakshagana, hindering initiatives such as training programs, research projects, and documentation efforts. Additionally, the absence of dedicated venues and infrastructure for Yakshagana performances limits its visibility and accessibility to a wider audience. In order to overcome these difficulties, our study created an automated system that correctly classifies Yakshagana characters according to their crown types and patterns of facial makeup using machine learning approaches. Several contributions that resulted from the proposed work are listed below:

• For character identification based on crown detection and facial makeup pattern analysis in Yakshagana pictures, this article explores deep learning techniques including YOLOv5 and Three Tier CNN.

• Approximately 10,000 images from the manually created Yakshachitra dataset served as the basis for the

[1] *Department of Master of Computer Applications, NMAM Institute of Technology, NITTE (Deemed to be University), India and Research Scholar, Institute of Computer Science and Information Science,, Srinivas University, India*
ORCID ID : 0000-0002-9009-7541
[2] *Dept. of Computer Science and Engineering, Srinivas University, Institute of Engineering and Technology, India,*
ORCID ID : 0000-0002-3957-1711
[3] *Institute of Engineering and Technology, Srinivas University, Institute of Engineering and Technology, India,*
ORCID ID : 0009-0000-8085-0491
[4] *Institute of Engineering and Technology, Srinivas University, Institute of Engineering and Technology, India,*
ORCID ID : 0009-0003-2958-7140
** Corresponding Author Email: anantham2004@gmail.com*

implementation of the proposed methods.

2. Related works

S.I No	Author(S)	Year	Methods	Dataset	Advantages
1	Chandan, G. et al. [1]	2018	SSD (Single Shot Detector)	Manually Created	SSD has been used for real-time object detection with enhanced accuracy, making it deployable in surveillance devices.
2	Ren, Y. et al. [3]	2018	Fast/Faster RCNN	VOC2007	Convolutional architectures are used in Fast or Faster RCNNs to detect objects. The VOC2007 dataset was used for the experiment, which shows that a classifier with a convolutional layer performs better at object detection.
3	Womg, A. et al. [4]	2018	Tiny-SSD	VOC207	The goal of the suggested approach is embedded object identification in real time. On the VOC2007 dataset, the model outperformed Tiny-YOLO despite being smaller. Its mean Average Precision (mAP) was 4.2% higher.
4	Sharma, V. et al. [5]	2019	Saliency driven Fast-er RCNN model	ASCAL VOC 2007, PASCAL VOC 2012 & CAMO_UOW	The experiment makes use of the AS-CAL VOC 2007, PASCAL VOC 2012, and CAMO_UOW datasets. Comparing the detection findings to traditional models, good results were found.
5	Xu, H. et al. [6]	2019	Reasoning-RCNN	COCO	The proposed method was successful in increasing the data on Vis-ualGenome by 16 percent, ADE by 37 percent, and COCO by 15 percent.
6	Mao, Q. C. et al. [7]	2019	Mini-YOLOv3	MS- COCO	The proposed Mini-YOLOv3, derived from YOLOv3, demonstrated superior accuracy compared to YOLOv3 with a detection time of 0.5 on the MS-COCO dataset, according to the authors' findings.

7	Zhu, Y. et al. [8]	2020	Tinaface based on Resnet	WIDERFACE	According to the viewpoint put forward, face detection and generic object detection are comparable. Compared to other state-of-the-art approaches, the ASFD-D6, a face detector method that is very efficient and effective, achieved an impressive 92.4% Average Precision (AP) on the WIDER FACE dataset.
8	Fang, Z., et al. [10]	2020	Resnet based face detector	FDDB, WIDERFACE	Accuracy and computational cost challenges in small object identification by proposing a system with optimized structure and loss functions. The experiment, conducted on FDDB and WIDERFACE, revealed that the proposed approach achieves a balanced mix of face detection accuracy and computational efficiency.
9	Al-ghanim, F. J. et al. [11]	2021	HOG and SVM Machine learning detector	DMFD	The proposed method aims to detect faces with makeup under diverse conditions. Through experimentation with the DMFD dataset using a HOG and SVM detector, the approach achieved a remarkable accuracy of 99.3% in identifying faces adorned with makeup.

The advantages of SSD object recognition over alternative methods were investigated by Anantha Murthy et al. in 2023 [9]. Based on how frequently an item category appears inside each box during prediction, the network, as seen on the feature map, adjusts the default boxes to better represent the form of the items. This results in scores. The output is mixed to help with the handling of objects of different sizes by utilizing multiple feature maps with

different resolutions. SSD is also used in Yakshagana photographs to identify the crown. [10] Rathgeb, C. et al. conducted experiments on the MIFS database to evaluate the proposed model, which effectively differentiates between makeup presentation attacks and genuine authentication attempts. Choi et al. [15] introduced Gaussian YOLOv3, revealing a significant improvement in mAP by 3.1 and 3.5 on BDD and KITTI datasets compared to YOLOv3. The proposed algorithm is recommended for achieving higher accuracy in autonomous driving applications. Yi et al. [16] presented a Tiny-YOLOv3 using K-means clustering

for pedestrian detection, resulting in higher accuracy for VOC2007 images. However, challenges were noted in detecting smaller objects with increased precision. Ou et al. [17] suggested using encoding and decoding to create a Moving Object Detection Method using ResNet-18. While it demonstrated superior performance compared to conventional algorithms, challenges were encountered in detecting smaller objects. Haque et al. [18] introduced a

ResNet Network with VGG for accurate object detection and image classification approach, which used ResNet for small item detection, demonstrated good accuracy in large-scale picture classification despite challenges with localization and training. A hybrid network using ResNet and YOLO combinations was proposed by Lu et al. [19] to detect numerous items in natural scene photos with an accuracy of 75.36%. A ResNet-based SSD model was published by Lu et al. [20] with the goal of outperforming VGG in terms of accuracy by 17.4% while using more computing power. In real-time target detection on the PASCAL VOC dataset, Bai et al. [21] paired a Residual Convolutional Neural Network with the YOLOv3-Tiny algorithm, obtaining 64.46% accuracy. Janahiraman et al. [22] compared SSD MobileNet V2 and Faster-RCNN for traffic light detection, revealing Faster RCNN's superiority with an accuracy of 97.01%. Fang et al. [23] proposed Tinier YOLO, four times smaller than Tiny-YOLOv3, to enhance detection accuracy. Yundong et al. [24] introduced Multiblock SSD to address limitations in detecting smaller objects. The model achieved a 96.6% accuracy on railway scene samples, surpassing traditional SSD by 9.2%.Anantha Murthy, et al., [25] conducted experiments in which they successfully identified disguises based on the types of crowns utilized. For this research they have used YOLOV5 and Three tire CNN and found that YOLOv5 is very much suitable for the disguise classification. As interpolated themes, folk, historical, imaginative, social, local temple legends, and awareness-oriented issues, these experiments are examined in the current conceptual piece. It examines the critical reactions that these projects have gotten from the general audience, artists, and scholars. The purpose of this study is to determine characters from the performers' specific crowns and facial makeup patterns.

3. Methodology

Using machine learning techniques, the proposed methodology focuses on character identification in Yakshagana images based on crown types and facial cosmetics patterns. Data collection, pre-processing, feature extraction, and classification are all part of this methodology. Figure 1 displays the block diagram of proposed model.

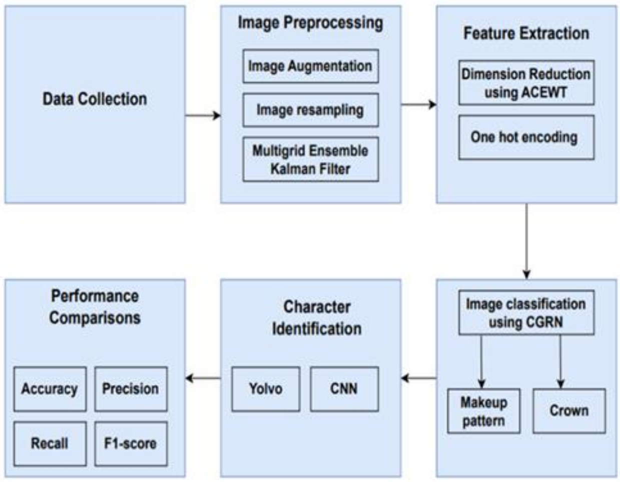

Fig 1: Block diagram of character identification

3.1 Character identification based on crown types and facial makeup pattern Yakshagana actors choose their roles in a predetermined order. To help with their identification and facial makeup, each figure has a unique crown. Character identification is based on the makeup patterns the performer uses and the many crowns the artist uses. The Tenkutittu Yakshagana features about ten different facial makeup patterns and about fifteen different types of crowns. The particular crown the performer wears and the facial makeup style they choose dictate the character option. For example, the Devi character wears the DeviMudi exclusively with a chandrakruthi as the pattern on the forehead, but the Vishnu characters have either pakdi, turai Kireeta, or Devi mudi on their foreheads. Mahisha's persona can be recognized by simply going with Mahisha Kireeta is sufficient. Anantha Murthy et al. provide a comprehensive explanation of the many crown kinds that performers use [22].

3.1.1 Data Collection

The hand collected dataset contains about ten thousand unique photos. Images of Yakshaganas with a variety of patterns were assembled, along with objects that were commonly seen during Yakshagana performances. A ratio of about 80:20 has been considered for the training and testing of the data.

3.1.2 Image Preprocessing

During the dataset generation phase, augmentation techniques such as rotation range = 30, zoom range = 0.2 and brightness between 0.6 and 0.7 were implemented. To alter the number of pixels within the photos, resampling of the photographs has been done. The images undergo preprocessing to remove noise from the input dataset. The Multigrid Ensemble Kalman Filter (MEKF) is utilized for this purpose, employing a sequential data assimilation method. The classic Ensemble Kalman Filter (EnKF) is extended or modified to include multigrid techniques in the Multigrid Ensemble Kalman Filter (MEKF). It increases the effectiveness and precision of state estimation in dynamic systems by combining the benefits of multigrid tactics and ensemble approaches. Table 2 gives the details of hyperparameter values of MEKF.

Table 2: Hyperparameter of Multigrid Ensemble Kalman Filter

Hyperparameter	Values
ensemble_size	10
inflation_factor	1
multigrid_levels	3
smoothing_factor	0.3
localization	0.1

3.1.3 Feature Extraction

Feature extraction is a method for reducing an image's dimensionality, or "dimensions," into manageable chunks. In order to achieve the necessary dimensions, the adaptive with concise empirical wavelet transform (ACEWT) is used to minimize image characteristics. After the feature extraction process, one-hot encoding is applied to the class labels rather than the image itself. This encoding technique transforms categorical labels into a binary matrix format, where each class is represented by a unique binary code. This ensures that the cyclic gate recurrent neural network (Cyclic-GRU) can effectively interpret and learn from the categorical nature of the class labels during the training process, enhancing its ability to make accurate predictions. Table 3 gives the details of hyperparameter values of ACEWT.

Table 3: Hyperparameter of Adaptive with concise empirical wavelet transform

Hyperparameter	Values
adaptation rate	0.02
adaptive control parameters	0.4
wavelet transform parameters	3
threshold value	0.2

3.1.4 Image Classification

A unique neural network called the Cyclic Gate Recurrent Neural Network (CGRNN) is utilized to identify characters such as Devi, Vishnu, Sarpa, or Mahisha. An expansion of the conventional Recurrent Neural Network (RNN) design is the Cyclic Gate Recurrent Neural Network (CycG-RNN). In order to boost the model's performance on tasks requiring comprehension of sequential patterns and to improve its capacity to capture long-range dependencies, it incorporates a cyclic gating mechanism.

Leveraging a Cyclic Gate Recurrent Neural Network (CGRU) architecture, we aim to capture temporal dependencies in performer images, allowing us to discern intricate details such as crowns and makeup patterns. The intricate cyclic gating mechanism in the CGRU facilitates the modeling of sequential patterns within a single image. By combining object detection techniques for locating crowns and makeup patterns with a classification model, we can determine the type of crown or makeup style present in each image. Ultimately, this comprehensive identification process serves as a foundation for predicting the character portrayed by the performer, providing valuable insights for character recognition in the context of diverse and dynamic visual appearances.

3.1.5 Character Identification

We used a two-step approach to character identification for Yakshagana actors, starting with identifying the characteristic crowns and elaborate makeup designs they sport. We tackle the important topic of character determination using YOLO (You Only Look Once) and the Convolutional Neural Networks (CNN) model. After a great deal of testing, we have found that the YOLO model regularly performs better than the CNN model, offering a higher level of character recognition accuracy. Because of its effectiveness in real-time object identification and capacity to manage several object classes at once, the YOLO model is positioned as a reliable option for identifying the subtle characteristics of Yakshagana performers, allowing for a more accurate representation of the characters they play. Table 4 gives the details of hyperparameter of YOLO model.

Table 4: Hyperparameter of YOLO Model

Hyperparameter	Values
batch_size	64
learning_rate	0.001
epochs	100
obj_thresh	0.5
nms_thresh	0.75
backbone	ResNet-50

3.1.6 Performance Comparison

Metrics including accuracy, precision, and F1 score were used to compare and evaluate the two models' performance, as well as their capacity to generalize to new patterns and objects outside of the training set. In terms of object detection and pattern recognition, YOLOv5 notably outperformed existing approaches with respect to classification accuracy when compared to the Three-tier CNN. Table 5 gives the performance analysis of the proposed models.

Table 5: Performance comparisons of algorithm

Algorithm	Category of Class	Precision	Accuracy	Recall	F1-Score
Three Tier CNN	Devi	0.906	0.90	0.906	0.90
	Vishnu	0.898	0.91	0.90	0.90
	Sarpa	0.904	0.90	0.91	0.91
	Mahisha	0.900	0.896	0.90	0.90
YOLOv5	Devi	0.957	0.982	0.95	0.941
	Vishnu	0.947	0.974	0.96	0.937
	Sarpa	0.958	0.980	0.93	0.959
	Mahisha	0.952	0.903	0.946	0.952

4. RESULTS AND DISCUSSION

The outcomes of a character identification experiment are shown in this section. The following diagrams show the Precision-Confidence, F1-Confidence, and Recall-Confidence curves. Experimentation with the yakshagana images depicted in figures 2-3 yields a state-of-the-art result in distinguishing between three separate classes of characters (Devi, Vishnu, and Sarpa). Four classes are combined to create the categories on the graph for each class. In Figure 3, it is evident that three features are taken into account for predicting the performer's character. These features encompass the facial skin color of the performer, with the recognition of character-specific facial attributes such as the necessity of a blue facial skin color for characters like Vishnu being particularly crucial for accurate identification. Figure 5 and 6 shows the analysis of the performance metrics such as recall, precision and f1-score along with the accuracy of the model. Figure 4 shows that facial makeup pattern does not need to be known to predict the Mahisha character, only crown identification is needed. Figure 7- 9 shows the confidence v/s the precision, accuracy and recall by including the four classes Devi, Vishnu, Sarpa and Mahisha.

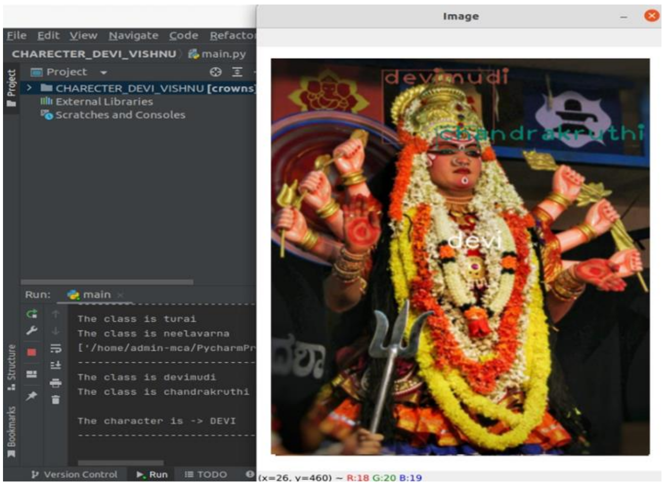

Fig. 2. Devi character identification using CNN

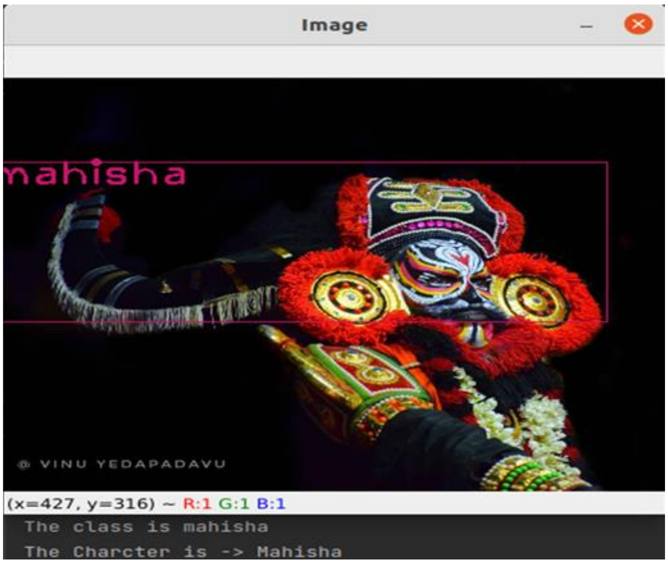

Fig. 3. Vishnu character identification using YOLOv5

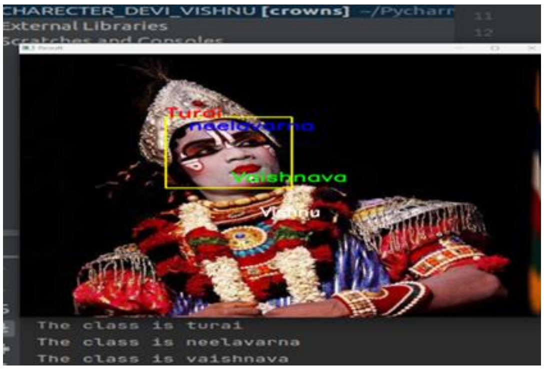

Fig. 4. character identification using YOLOv5

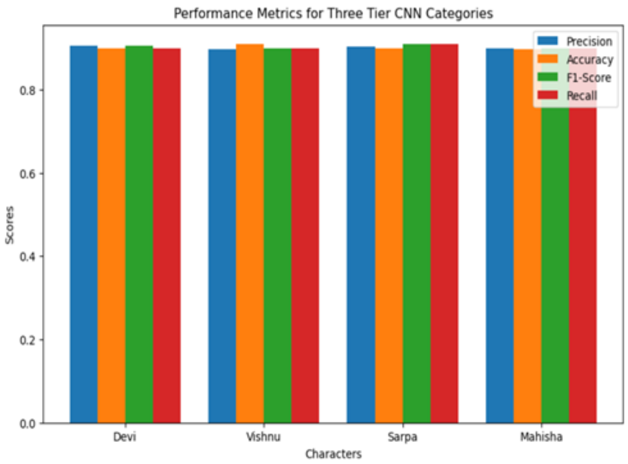

Fig 5: Performance metrics of three tier CNN

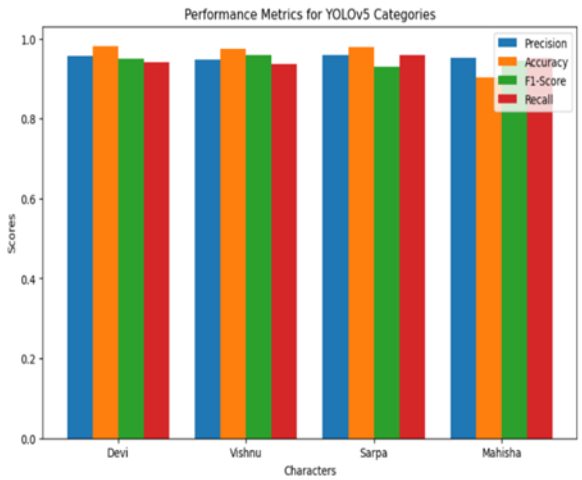

Fig 6: Performance metrics of three tier YOLOv5

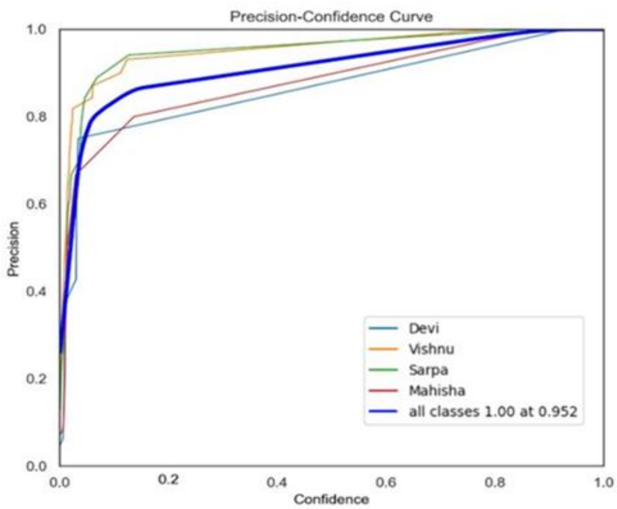

Fig 7: Confidence v/s Precision Curve

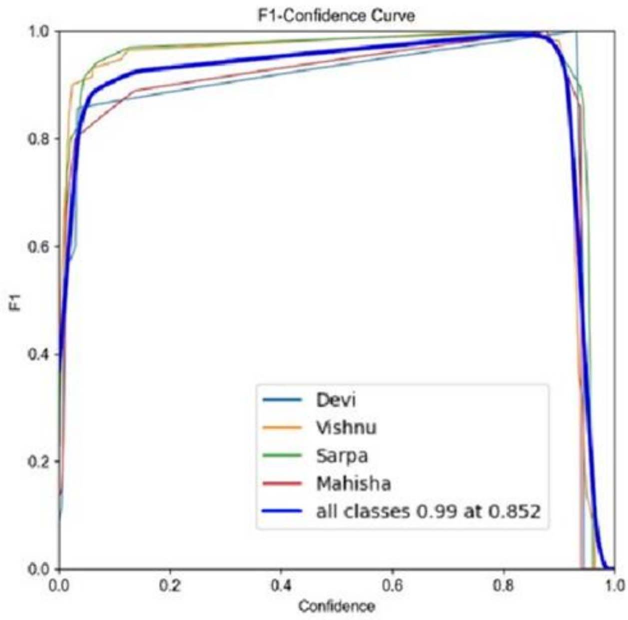

Fig 8: Confidence v/s F1-score Curve

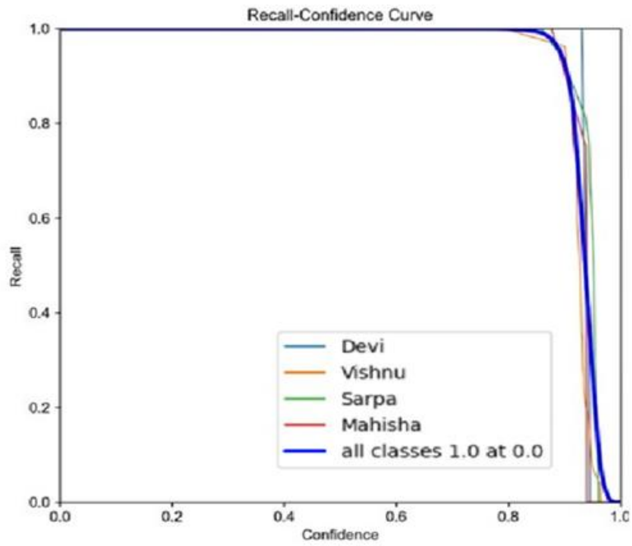

Fig 9: Recall v/s F1-score Curve

5. Conclusion and Future work

Character identification in Yakshagana images mostly depends on crown and cosmetics pattern detection. On the basis of crown type recognition, machine learning approaches like YOLOv5 and Three-tier CNN have demonstrated good results in disguise categorization. YOLOv5, in particular, outperforms the Three-tier CNN model with an astounding 95% accuracy rate. This accomplishment highlights the advantages of every strategy in its own field. YOLOv5 is distinguished by its ability to draw conclusions quickly, accurately identify several crown items in a single image, and have a simple design that makes training and development easy. Its ability to discriminate between different disguises in Yakshagana depictions is greatly enhanced by its skill in identifying facial traits and crown attributes. However, variables like dataset size, should be considered while choosing a machine learning method and architecture.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Chandan, G., Jain, A., & Jain, H. (2018, July). Real time object detection and tracking using Deep Learning and OpenCV. In 2018 international conference on inventive research in computing applications (ICIRCA), 1(1), 1305-1308.

[2] Murthy, A., Devadiga, P. and Nethravathi, P.S., FASTER R-CNN APPROACH FOR DETECTING SMALLER JEWELS IN YAKSHAGANA IMAGE.

[3] Ren, Y., Zhu, C., & Xiao, S. (2018). Object detection based on fast/faster RCNN employing fully convolutional architectures. Mathematical Problems in Engineering, 2018, 1-7.

[4] Womg, A., Shafiee, M. J., Li, F., & Chwyl, B. (2018, May). Tiny SSD: A tiny single-shot detection deep convolutional neural network for real-time embedded object detection. In 2018 15th Conference on Computer and Robot Vision (CRV), 1(1), 95-101.

[5] Sharma, V., & Mir, R. N. (2019). Saliency guided faster-RCNN (SGFr-RCNN) model for object detection and recognition. Journal of King Saud University-Computer and Information Sciences,1(1),1-12.

[6] Xu, H., Jiang, C., Liang, X., Lin, L., & Li, Z. (2019). Reasoning-rcnn: Unifying adaptive global reasoning into large-scale object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1(1), 6419-6428.

[7] Mao, Q. C., Sun, H. M., Liu, Y. B., & Jia, R. S. (2019). Mini-YOLOv3: real-time object detector for embedded applications. IEEE Access, 7(1), 133529-133538.

[8] Zhu, Y., Cai, H., Zhang, S., Wang, C., & Xiong, Y. (2020). Tinaface: Strong but simple baseline for face detection. arXiv preprint arXiv:2011.13183.,1(1), 1-9.

[9] Murthy, A., Shetty, P.P. and Nethravathi, P.S., CROWN DETECTION IN YAKSHAGANA IMAGES USING SSD.

[10] Fang, Z., Ren, J., Marshall, S., Zhao, H., Wang, Z., Huang, K., & Xiao, B. (2020). Triple loss for hard face detection. Neurocomputing, 398, 1(1), 20-30.

[11] Zhang, H., Sun, M., Li, Q., Liu, L., Liu, M. and Ji, Y., 2021. An empirical study of multi-scale object detection in high resolution UAV images. Neurocomputing, 421, pp.173-182.

[12] Murthy, A., Nayak, N. and Nethravathi, P.S., UNRAVELLING THE MYSTIQUE OF CROWNS IN YAKSHAGANA IMAGES THROUGH ALGORITHMIC DETECTION.

[13] Al-ghanim, F. J., & mohsin Al-juboori, A. (2021). Detection of Human Faces Covered with disguise and Makeup. Journal of Al-Qadisiyah for computer science and mathematics, 13(3), 68.

[14] Rathgeb, C., Drozdowski, P., & Busch, C. (2021, January). Detection of makeup presentation attacks based on deep face representations. International Conference on Pattern Recognition (ICPR), 1(1), 3443-3450.

[15] Choi, J., Chun, D., Kim, H., & Lee, H. J. (2019). Gaussian yolov3: An accurate and fast object detector using localization uncertainty for autonomous driving. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 1(1), 502-511.

[16] Yi, Z., Yongliang, S., & Jun, Z. (2019). An improved tiny-yolov3 pedestrian detection algorithm. Optik, 183, 17- 23.

[17] Ou, X., Yan, P., Zhang, Y., Tu, B., Zhang, G., Wu, J., & Li, W. (2019). Moving object detection method via ResNet-18 with encoder–decoder structure in complex scenes. IEEE Access, 7(1), 108152-108160.

[18] Haque, M. F., Lim, H. Y., & Kang, D. S. (2019, January). Object detection based on vgg with resnet network. International Conference on Electronics, Information, and Communication (ICEIC), 1(1), 1-3.

[19] Lu, Z., Lu, J., Ge, Q., & Zhan, T. (2019, July). Multi-object detection method based on YOLO and resNet hybrid networks. In 2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM), 1(1), 827-832.

[20] Lu, X., Kang, X., Nishide, S., & Ren, F. (2019, December). Object detection based on SSD-ResNet. International Conference on Cloud Computing and Intelligence Systems (CCIS), 1(1), 89-92.

[21] Bai, Z., & Jiang, D. (2019, November). On the multi-scale real-time object detection using resnet. In Chinese Conference on Pattern Recognition and Computer Vision (PRCV), 1(1), 63-73.

[22] Janahiraman, T. V., & Subuhan, M. S. M. (2019, October). Traffic light detection using TensorFlow object detection framework. In 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET), 1(1), 108-113.

[23] Fang, W., Wang, L., & Ren, P. (2019). Tinier-YOLO: A real-time object detection method for constrained environments. IEEE Access, 8(1), 1935-1944.

[24] Yundong, L. I., Han, D. O. N. G., Hongguang, L. I., Zhang, X., Zhang, B., & Zhifeng, X. I. A. O. (2020). Multi-block SSD based on small object detection for UAV railway scene surveillance. Chinese Journal of Aeronautics, 33(6), 1747-1755.

[25] Murthy, A. ., & Kulkarni, S. . (2024). Crown-Centric Disguise Classification In Yakshagana Imagery Using Deep Learning Techniques. Migration Letters, 21(S4), 914–929.