Scope
Volume 14 Number 1 March 2024

Enhanced Modeling by Unveil the Feature for Learning Celebrity Cartoon Faces

Prajna S[1], D S Guru[1], Shivaprasad D L[1], and Vinay Kumar N[2]
[1], Department of Studies in Computer Science, Manasagangotri, University of Mysore, Mysuru 570009
[2]Standard chartered Bank-GBS, Bangalore, India.

Abstract. In this paper, we propose a new approach that aims to uncover features that can assist in learning celebrity cartoon faces. To recognize cartoon faces, we have tailored the FaceNet architecture. The extracted features are then learned using both a conventional learning model and a convolution model. Furthermore, the Chi-score method is employed to achieve feature reduction to have an efficient yet effective classification. To demonstrate the effectiveness of this approach, we conducted extensive experiments on the Cartoon Faces in the Wild (IIIT-CFW) celebrity cartoon face database, which contains 100 distinct celebrities. In comparison to the existing database, we created 50 different categories namely UOM-Dataset, and examined its performance. The results of the experimentations reveal that the proposed method outperforms several other existing methods, including the state-of-the-art method for celebrity cartoon face recognition.

Keywords: Cartoon Images, Deep Features, Support Vector Machine, Cartoon Face Recognition.

1 Introduction

In our day-to-day life, cartoons have become a prominent and prosperous part of entertainment. Cartoons have a wide range of applications in entertainment, education, advertisement, and communication, such as Newspapers, instant messaging, Etc., by mocking or spoofing the contextual and situational interpretations of the real-world scenario. Thus, the extensive usage of cartoons has drawn research attention in the field of computer graphics and multimedia. Designing cartoons relies upon cartoonists' perception, and there existsa significant gap between modeling real faces and cartoon faces. Hence, the approaches used to detect and recognize a real face may not be feasible for cartoons. Cartoon face recognition [1] is the primary step in the direction of diverse applications, including cartoon face verification, cartoon image retrieval [2],gender prediction, age estimation, cartoon face synthesis, and photo-to-cartoon search. Though the problem of celebrity cartoon face recognition is closely related to real face recognition [3], it poses many additional challenges due to their artistic variations and high caricatures.

Figure. 1. Illustration of celebrity cartoon images exhibiting high intra-class variations (A) celebrity cartoon images of Class Amir khan (b) celebrity cartoon images of Class Abraham Lincoln (C) celebrity cartoon images of Class Aishwarya rai (D) celebrity cartoon images of Class Lucille Ball

Figure 1 illustrates a few pictorial samples of celebrity cartoon images, they exhibit high intra-class variation among the samples of celebrity cartoon images; indeed, it is difficult to categorize the samples without any ambiguity. Different models have been proposed which differ in the features, representation scheme adopted, dataset and the classifiers adopted. In spite of existing models, finding an optimal set of discriminating features and also deciding upon the best classifier for learning celebrity cartoon images are still open issues.

In this direction, regarding the recognition of cartoon faces, we found a couple of works related to cartoon character recognition. In this work [8], face detection of cartoon characters was explored using primitive features such as skin color regions and edges extracted from the given cartoon character images. X. Gao et al. [9] explored the image retargeting techniques for mobile comics by adopting the comic content to mobile screens for better visualization. The works carried out were in concern with cartoon characters. Over the recent years, many works of literature aim to automatically generate a cartoon-style face with an input RGB image [10], [11] or sketches [12], [13], while cartoon face recognition has been less explored and remains a challenging problem. Automatic cartoon face recognition performs poorly in daily applications due to the lack of large-scale real-world datasets. Existing datasets (Web Caricature [14], IIIT-CFW [15], Manga109 [16]) contain limited examples and fail to reflect the accurate distribution of the cartoon faces. Among the existing methods, Takayama et al. [8] extracted features that reflect skin color, hair color, and hair quantity for cartoon face recognition. Saito et al. [17] proposed a CNN-based model for cartoon image retrieval. Zhou et al. [18] constructed a ToonNet that contains thousands of cartoon-styled images and introduced several techniques for building a deep neural network for cartoon face recognition. Zheng et al. [19] proposed a meta-continual learning method capable of Jointly learning from heterogeneous modalities such as sketch, cartoon, and caricature images. These pioneering methods bring inspiration for the research on cartoon face recognition. In this work, we study automatic cartoon face recognition by adopting the iCartoonFace [20] and Danbooru [21] datasets, the largest and most comprehensive datasets for cartoon face recognition. The shape characteristics can be better exploited for accurate cartoon face recognition.

The entire literature survey has been carried out towards cartoon character recognition but has yet to be dominated by celebrity cartoon face recognition. However, we found this work [22], the Multitask Cascaded Convolutional Network (MTCNN) architecture is used for face detection and has been compared with conventional methods. Two different approaches for recognition were recommended; one based on transfer learning by combining the feature learning capability of the Inception v3 network with the feature recognizing capability of SVM, and the other one, is a Hybrid Convolutional Neural Network (HCNN) framework trained over a fusion of pixel values and 15 manually located facial key points. These methods are evaluated on the (IIIT-CFW) database. The survey concludes with an understanding that all works rely on cartoon character recognition than celebrity cartoon face recognition due to the inadequacy of the work carried out related to celebrity cartoon face recognition.

In the current work, our intuition is to consider the problem of celebrity cartoon face recognition as it is closely related to real face image recognition based on enhanced modeling by uncovering the feature for learning celebrity cartoon faces.

Considering all these factors in this work, we recommend the FaceNet architecture for feature extraction using celebrity cartoon images. The recommended model also performs well when compared with existing contemporary models specifically designed for recognizing celebrity cartoon faces. We compared the recommended model to other well-known learning models, such as Random Forest [3], Gaussian Naive Bayes [4], K-Nearest Neighbor [5], Decision Tree [6], and SoftMax classifier [7] on the IIIT-CFW dataset. Our extensive experimental study reveals that the model performs better when FaceNet is integrated with SVM than other classical supervised learning algorithms. Furthermore, by comparing the chi-scores of the features according to relevance, the number of features is reduced. Also, we have created a supporting database and compared the recommended model to that and against the existing model.

Overall, significant contributions of this work are,

- Investigated appropriate classifiers for recognition with varying hyperparameter tuning.
- Conducted a quantitative analysis to determine the optimal number of features that enhance the performance of the selected classifier.
- Created a Dataset of our own 50 Distinct Celebrity Cartoon Faces for the study.
- Comparative study of the proposed model against the state-of-the-art models.

The rest of the paper is structured as follows: In Section 2, the proposed models are presented. In Section 4 the details of experimentations along with results are summarized. A comparative study is presented in Section 5. Finally, Section 6 follows conclusion and future works are presented.

2 Proposed Model

The proposed model includes mainly three stages, preprocessing, deep feature extraction, and classification using the SVM learning model. The typical architecture of the recommended model is given in Figure 3.

2.1 Preprocessing

In this work, the pre-processing technique is attained by detecting the region of interest (celebrity cartoon face) in a given cartoon image and cropping the detected image concerned with the area of interest. However, automation of the detection of celebrity cartoon faces in cartoon images is kept beyond the scope of this paper as it is our future target.

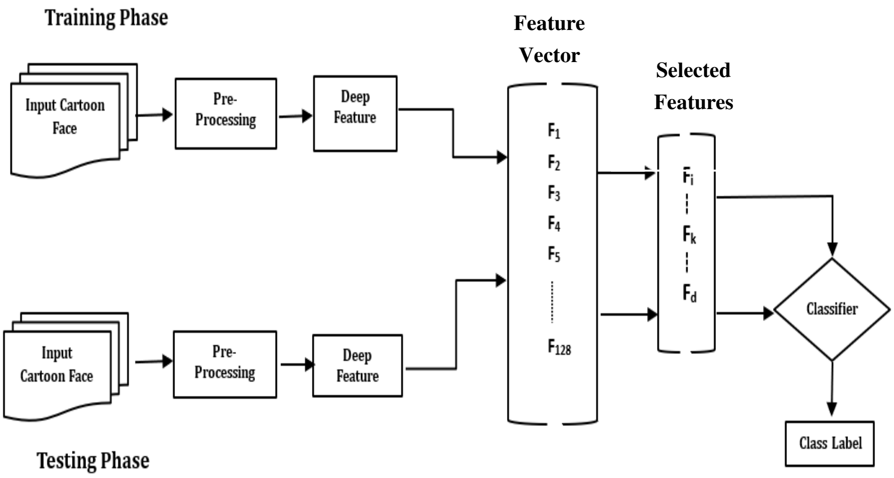


Figure. 2. General Architecture of the Proposed Model.

2.2 Feature Extraction

The detected celebrity cartoon face images are used to extract the face encodings. FaceNet is a well-known architecture for face recognition. It is a pre-trained architecture for real-face images showing a significant margin of performance. This motivated us to adopt the FaceNet architecture [23] for celebrity cartoon face recognition. FaceNet maps each image into a compact Euclidean space such that the distance corresponds to face similarity, and the compact embedding is generated using a triplet-based loss function and triplet selection based on the extensive margin nearest neighbor. The main characteristic between FaceNet and other techniques is that it learns the mapping from images and creates unique embedding rather than using any bottleneck layer for recognition.

2.3 Feature Selection

All the extracted features may not contribute towards learning. When it comes to selecting features, the filter-based method is often preferred over the wrapper-based method because it can handle large datasets more efficiently. This method selects features independently of a learning algorithm, making it easier for us to rank and consider them for learning purposes. Hence, we explored a Filter-based feature selection technique to select the subset of features from the original set of features through feature ranking criteria namely chi-square. Each feature is ranked based on its relevancy. Further, the Chi-square feature selection aims to select features highly dependent on the Target. [31]. All the considerable individual features are ranked and sorted based on relevancy, and the top-ranked features are selected and considered for learning. Whereas a subset of k features is chosen such that the ratio of the sum of the scores of k features in the subset to the sum of scores of all features is more significant than a threshold. The threshold is a confidence score in selecting features as given in equation (1). The performance is studied using these ranked features by sequentially introducing a feature-by-feature for learning up to the confidence score.

$$min\, k \left(\sum_{i=1}^{k} Si \,\middle|\, \sum_{i=1}^{k} Si \right) > CONFIDENCE\_SCORE \qquad (1)$$

3 Experimentation

3.1 Celebrity Cartoon Face Database

In this section, we will discuss the dataset that we used for our experimentation. We utilized Cartoon Faces in the Wild (IIIT-CFW), which is a standard dataset specifically created for celebrity cartoon face recognition. Figure 3(a-b) shows a few sample images from the dataset. This dataset contains 8,928 celebrity cartoon faces of 100 popular personalities. The number of samples in each category ranges from 11 to 299. In this section, however, it should be noted that the IIIT-CFW database contains not only pure cartoon images but also sketches in a few classes.

Figure. 3. Pictorial Samples of IIIT-CFW dataset (a) Samples of class Amir Khan (b) Samples of class Indira Gandhi.

3.2 UoM Celebrity Cartoon Face Database

We created our dataset, UoM-Cartoon Faces, which includes 50 distinct classes with 30 samples. Unlike IIIT-CFW datasets, UoM-Cartoon Faces comprises exclusively cartoon images. Demonstrated in Figure

4(c-d)depicts a few sample images from the dataset. The samples across the class are balanced .

Figure. 4. Pictorial Samples of UoM Dataset (c) samples of class Narendra Modi (d) Samples of class Mayavati

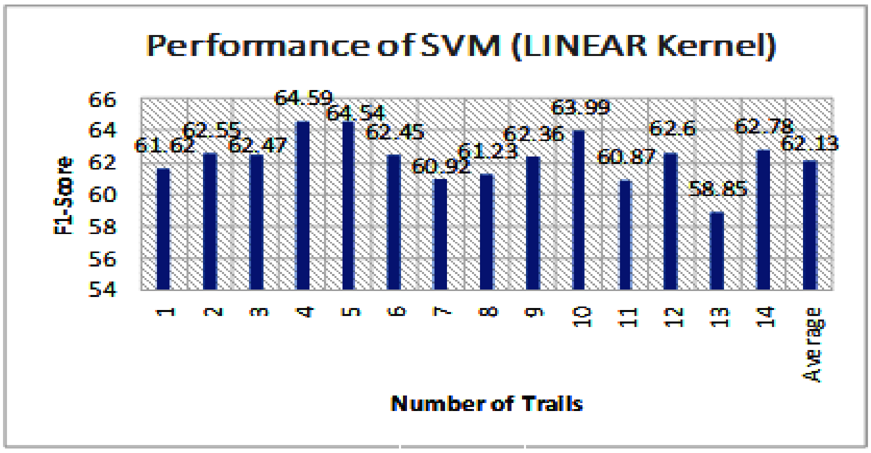
4. Experimental Setup and Results

In this section, we present the details of the experimentation conducted during the research and the obtained results.

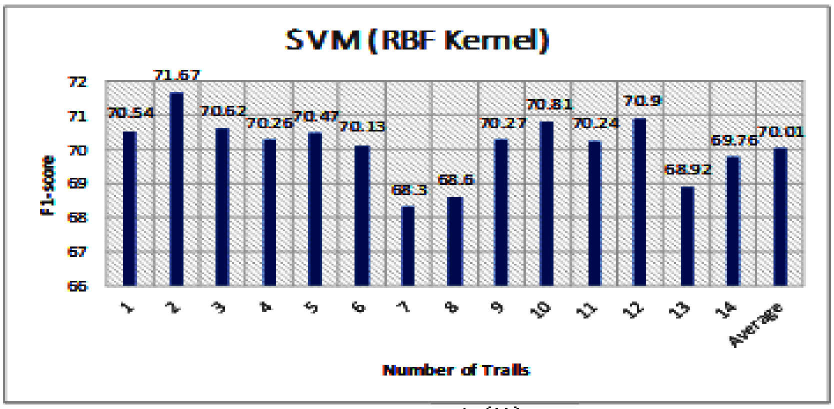
We have used the two sets of databases for our experimentations:
1. IIIT-CFW dataset The Cartoon Faces
2. UoM – Cartoon Face dataset

Initially, we considered the IIIT-CFW dataset and the FaceNet architecture, which is recommended for real faces, is adapted for extracting features. We used this pre-trained architecture to extract '128-dimensional embeddings. We randomly selected 10 classes out of 100 and divided the chosen class samples into two phases - training and testing in a ratio of 80:20 and then tested the performance of our proposed model using conventional classifiers such as DT[6], RF [27], KNN [28], GNB [30], and SVM [32] with varied hyperparameter tuning, as well as fully connected deep architecture such as FaceNet and CNN are explored. To ensure accurate results multiple trials (T=15) are conducted and performance of the model is evaluated using a well-known F-Measure metric [32]. It considers both precision and recall metrics which is suitable for analyzing imbalanced [30] data, rather than considering the accuracy as a metric. Furthermore, the features in the feature space are analyzed to reduce their dimension by a filter-based feature selection technique using Chi-square score as explained in section 3.2.

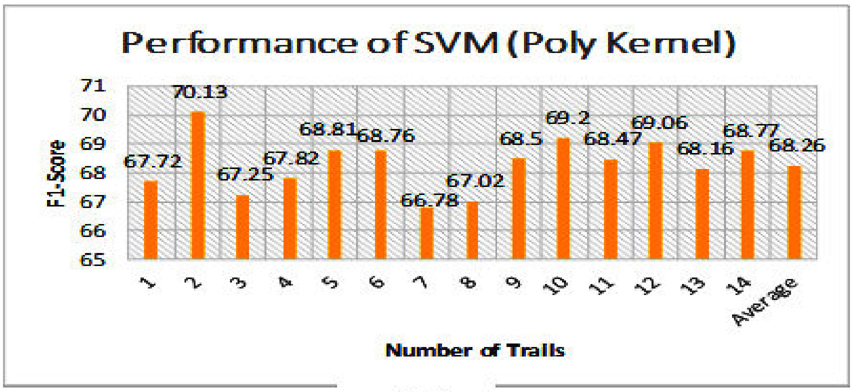
Experimental results on IIIT-CFW Dataset. Using IIIT-CFW dataset this dataset experimentation is carried out and results are presented in this section. In Figure 5 shows the performance of the SVM classifier with distinct kernels. It is noticeable that SVM with RBF kernel outperforms the other, with the average F1-score across all the kernels, under varying numbers of trials and hyperparameter tuning, which indeed helps us to recommend a learning model for celebrities in cartoon images with F1-score of 70.01 percent as state-of-art-of results for recognition. To evaluate the performance of a model it is essential to choose a good kernel function. Choosing the kernel function also emphasizes the kind of dataset. We here explore distinct kernels in our study.



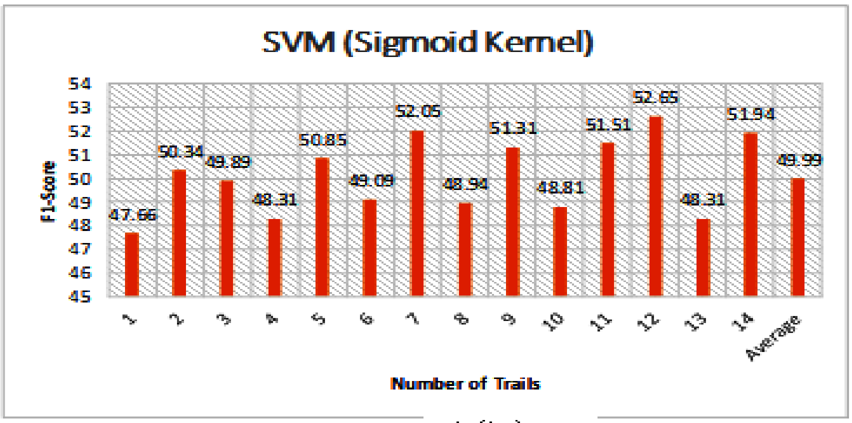
A (i)



A (ii)

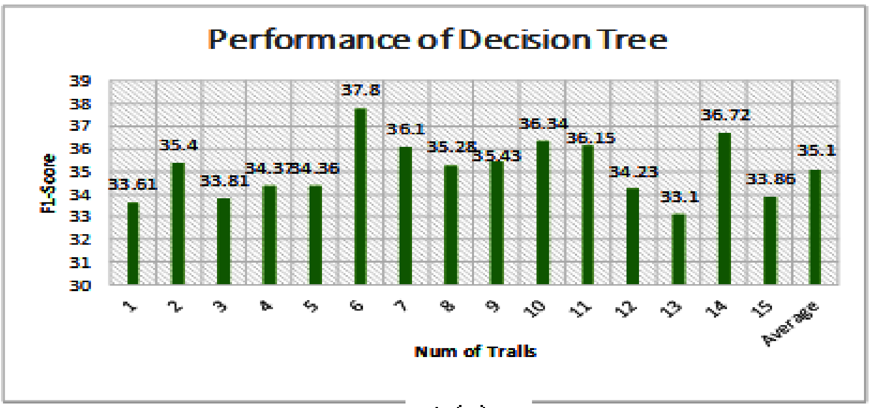


A (iii)

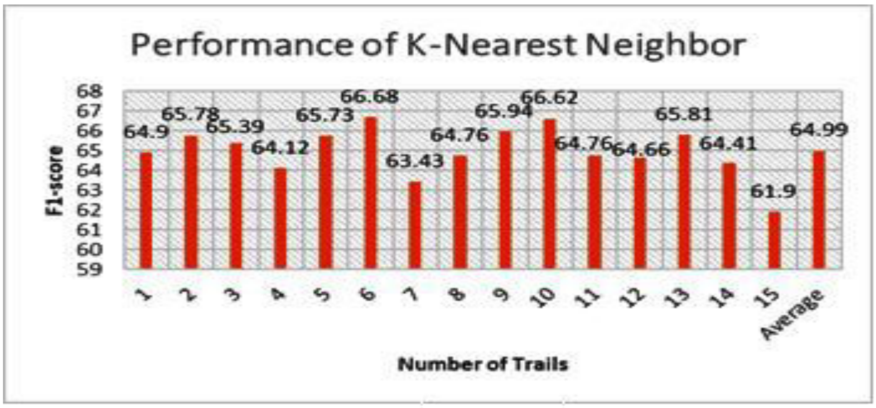


A (iv)

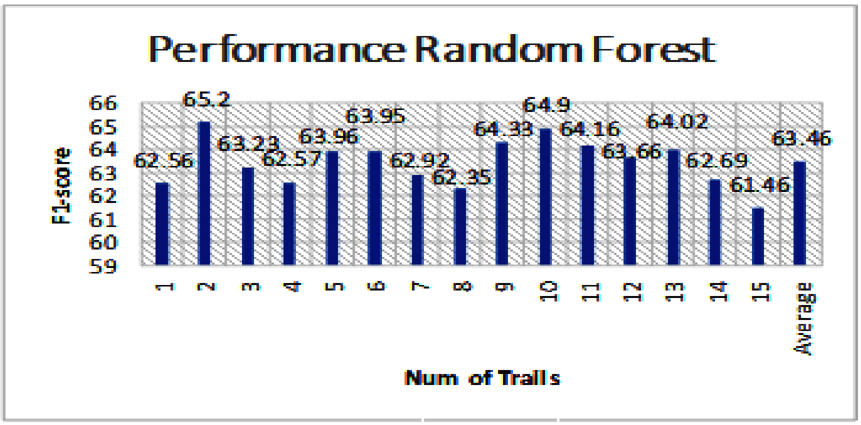
Fig. 5.A (i-iv) Performance of Support Vector Machine (SVM) with distinct kernels under varied number of trails.



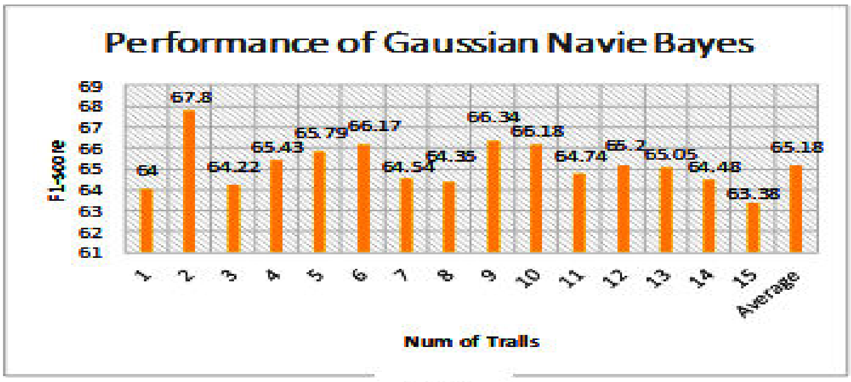
A (v)



A (vi)

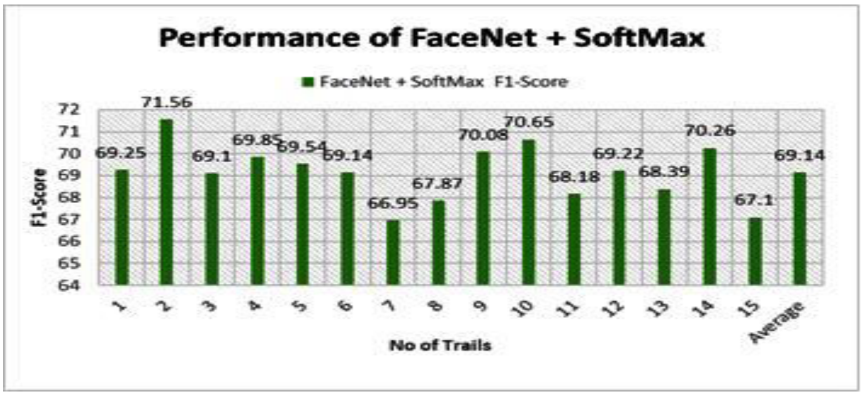


A (vii)

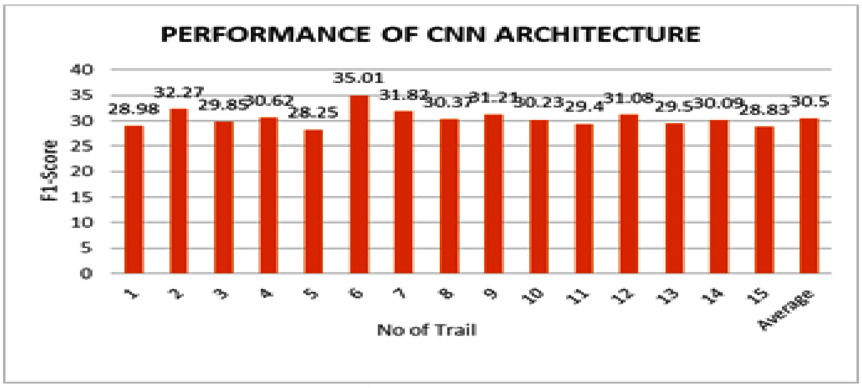


A (viii)

Fig. 6.A(v-viii) Performance of conventional classifiers under varied number of trails.



B (i)



B (ii)

Fig. 5.B Depicts the Performance of fully connected architecture.

Figure 5.A (i-iv) show the performances of SVM classifier with distinct kernels under varying trails. Experimentally it is evident that SVM with RBF kernel outperforms other conventional classifiers. Figure5.A(v-vii) depicts the performance of distinct classifiers under varied trails, Figure 5.B(i) Performance of FaceNet Fully connected architecture, Figure 5.B(ii) CNN Fully connected architecture are graphically addressed. It is observed that the performance of the fully connected architecture is not on par to our recommended model. However, the overall study shows that our recommended classifier performs better than the others. In addition, our focus is on analyzing the features in the feature space. It is accomplished through a Filter based Feature selection, namely Chi-square score based on feature ranking criterion, which evaluates each feature and assigns a rank based on relevancy. It is experimentally, evident that the performance deteriorates as Number of classes increases [33]. Table 1 depicts class-wise performance with and without feature selection. The study infers that our recommended model outperforms the existing one. However, it's important to note that we analyzed the performance on 100 classes of the IIIT-CFW dataset.

Table 1. Class-Wise Performance of SVM classifier with and without Feature Selection using IIIT-CFW Dataset.

Number of Classes	**10 Class**	**20 Class**	**30 class**	**40 Class**	**50 Class**	**60 Class**	**70 Class**	**80 Class**	**90 Class**	**100 class**
Without Feature Selection										
F1-Score (128-Feature)	83.11 (128)	83.69 (128)	78.94 (128)	76.97 (128)	74.82 (128)	74.00 (128)	72.87 (128)	72.57 (128)	71.05 (128)	70.71 (128)
With Feature Selection										
Number of Feature Selected attaining the State-of-the-art-result	46	77	72	83	114	99	105	124	101	118
F1-score	83.18	83.47	78.94	77.84	75.07	74.28	73.65	72.62	71.25	71.96

Number of Feature Selected Attaining Max F1-score	90	107	120	115	119	124	127	116	126	111
F1-Score	90.29	87.4	81.99	79.89	77.99	76.93	74.88	74.92	71.99	72.42

The table above displays the performance of our model with and without feature selection for different numbers of classes. Our model achieved an F1 score of 83.11% without feature selection when trained on 10 classes. We observed that by employing feature selection for only 46 features, the F1 score crossed 83.18%. The highest F1 score of 90.29% was achieved with 90 features. For 20 and 30 classes, our model scored 83.69% and 78.94% with 128 features. After feature selection, we achieved scores higher than those without feature selection, i.e., 83.47% and 78.94%, with only 77 and 72 features, respectively. The maximum F1-score for 20 and 30 classes was attained with 107 and 120 features, achieving scores of 87.4% and 81.99%, respectively. We noticed that the performance of our model decreased as the number of classes increased. Our model's performance without feature selection was 76.97% and 74.82% for 40 and 50 classes, respectively, with 128 features. However, after feature selection, we crossed the F1-score scores of 77.48% and 75.07% with only 83 and 114 features, respectively. The highest F1 score for 40 classes was 79.89% with 115 features, while for 50 classes, it was 77.99% with 119 features.

For 60 and 70 classes, our model's performance without feature selection was 74.00% and 72.87%, respectively, using 128 features. However, after employing feature selection, we achieved scores higher than those without feature selection, i.e., 74.28% and 73.65%, with only 99 and 105 features, respectively. The highest F1 score obtained for the 60-class dataset was 76.93% using 124 features. For the 70-class dataset, the highest F1-score we achieved was 74.88% using 127 features.

For 80 classes, our model's performance without feature selection was 72.57% using 128 features. After employing feature selection, we crossed an F1 score of 72.62% with only 124 features. The highest F1 score we achieved was 74.92% with 116 features. For 90 classes, our model's performance without feature selection was 71.05% using 128 features. With feature selection, we were able to cross an F1-score of 71.25% using only 101 features, while the maximum F1 score we achieved was 71.99% with 126 features.

Lastly, for 100 classes, our model's performance without feature selection was 70.71% using 128 features. However, after employing feature selection, we were able to cross the F1-score of 71.96% with only 118 features. The maximum F1 score obtained is 72.42 with 111 features only, which is lower than the state-of-the-art results [23] for recognizing celebrity cartoon faces.

5 Comparative Study

This section presents a comparative analysis of the proposed two models against themselves and also compares the best proposed model against the several other existing contemporary models.

5.1 Across the Proposed Models

i. A comparison of the recommended models is conducted to compare them against themselves. Our recommended model explores a suitable classifier for learning celebrity cartoon images. It is observed that the Support Vector Machine with RBF kernel outperforms other classifiers. On average, it achieves an impressive F1 score of 70.01 percent.

ii. We also compared the performance of our model with fully connected CNN, and FaceNet architectures on the same IIIT-CFW dataset. The results of the SoftMax classifier (see Figure B(i-ii)) are not satisfactory, which motivated us to learn conventional classifiers.

iii. Our recommended model performance is studied with and without feature selection, and it is found to be superior. We infer from Table 1. that we achieve 70.71 for 100 classes. After incorporating feature selection results obtained is an F1-score of 71.96 for 118 features and 72.42 for 111 features.

5.2 Across the Existing Models

In order to bring out the excellence of our recommended system when it is compared against to existing model.We conducted a comparison between our recommended system and a contemporary model to demonstrate its superiority. Jha et al. (2018) also conducted class-wise comparisons, is shown in Table 2. It is observed that the contemporary model achieved 84.81% for 20 classes, 67.49% for 50 classes, and 60.94% for 100 classes on balanced data. Our model achieved 85.11% accuracy for 20 classes, 76.37% for 50 classes, and 72.29% for 100 classes on imbalanced data. Finally, our model excels more than the contemporary model.

Table 2. Depicts overall SVM Classifier under varied number of classes without Feature Selection.

Accuracy Score				**Attempt of Balancing Data**
Models	**20 Class**	**50 Class**	**100 Class**	
Inception V3+SVM (Jha et. al., 2018)	84.81%	67.49%	60.94%	Yes
Proposed Model FaceNet+SVM	85.11%	76.37%	72.29%	No

Table 3. Performance comparison of recommended model with respect to other existing models on IIIT-(CFW) Dataset for recognition celebrity cartoon images over 100 classes.

<table>
<tr><th>Model</th><th colspan="2">F1-score</th><th>Attempt for Balancing Data</th></tr>
<tr><td>Deep Features + SVM(RBF)
Proposed Model</td><td colspan="2">70.0
(128 features)</td><td>No</td></tr>
<tr><td>WITH FEATURE SELECTION
F1-Score</td><td>(Minimum)
69.98
(99 features)</td><td>(Maximum)
72.42
(111 features)</td><td>No</td></tr>
<tr><td colspan="4">Contemporary Model</td></tr>
<tr><td>HCNN (Jha et. al., 2018)</td><td colspan="2">0.649</td><td>Yes</td></tr>
<tr><td>INCEPTION V3+SVMJha et. al., 2018)</td><td colspan="2">67.0
(2048 Features)</td><td>Yes</td></tr>
</table>

Table 3 demonstrates that our proposed model outperforms the contemporary model. Our model attained an F1 score of 70% by employing 128 features. By employing feature selection with 99 features, we achieved a minimum score of 69.98%, and with a maximum of 111 features, scores of 72.42%. On the other hand, the contemporary model scored 67% by using 2048 features. Nevertheless, our model is superior to the contemporary model with a score of 70% as the state-of-the-art.

We evaluated the performance of our model by testing it on our UOM dataset. This dataset includes 1500 cartoon images of 50 celebrities, with 30 images per class. The purpose of this dataset is to demonstrate the model's performance and highlight the significance of highly caricatured images. To underscore the creation of this dataset, we applied the recommended model to it and conducted a study.

Table 4. Depicts Performance of SVM Classifier of 10 and 20 classes respectively under varied number of trails with and without Feature Selection on UoM Dataset.

Number of Classes	**10class**	**20class**	**30class**	**40Class**	**50class**
F1-Score (128-Feature)	67.45 (128)	71.76 (128)	61.81 (128)	58.62 (128)	55.96 (128)
Number of Feature Selected attaining the State-of-the-art-result	80	68	66	84	69
F1-score	67.81	72.02	62.58	58.64	56.71
Number of Feature Selected Attaining Max F1-score	54	82	116	109	94
F1-Score	78.80	73.63	68.01	63.06	62.95

The following table shows the results of our proposed model on the UoM dataset. We evaluated the model's performance by analyzing it class-wise. For 10 classes, the model achieved a score of 67.45% using 128 features. However, after using feature selection, the score improved to 67.81% with only 80 features. The maximum score of 78.80% was achieved using just 54 features. For 20 classes, the model scored 71.76% using 128 features with feature selection. The score improved to 72.02% with only 68 features. The maximum score of 73.63% was achieved using just 82 features. For 30 and 40 classes, the model achieved scores of 61.81% and 58.62% respectively without feature selection. However, with feature selection, the scores improved to 62.58% with 66 features and 58.64% with 84 features. The maximum scores of 68.01% and 63.06% were achieved using 116 and 109 features respectively. For 50 classes, the model achieved a score of 55.96% without feature selection. However, with feature selection, the score improved to 56.71% using 69 features. The maximum F1-score of 62.65% was achieved with 94 features.

Based on our analysis, we conclude that the UoM dataset's highly caricatured images resulted in lower performance compared to the IIIT-CFW dataset.

Table 5. Depicts Performance of SVM Classifier under varied number of classes with and without Feature Selection.

Dataset		**10 Class**	**20 Class**	**30Class**	**40Class**	**50Class**
UOM (Balanced data)	Without Feature Selection F1-Score	67.45 (128)	71.76 (128)	61.81 (128)	58.62 (128)	55.96 (128)
	With Feature Selection F1-Score	78.8 (54)	72.02 (68)	67.57 (117)	62.03 (90)	60.45 (94)
	With Feature Selection Maximum F1-Score	78.8 (54)	73.63 (127)	68.1 (116)	62.38 (66)	62.65 (94)
IIIT- Dataset (Un-Balanced data)	Without Feature Selection F1-Score	83.11 (128)	83.69 (128)	78.94 (128)	76.97 (128)	74.82 (128)
	With Feature Selection F1-Score	88.62 (62)	85.96 (88)	80.81 (86)	78.47 (85)	76.42 (99)
	With Feature Selection Maximum F1-Score	90.29 (90)	87.4 (107)	81.99 (120)	79.89 (115)	77.99 (119)

We compared the performance of the proposed approach with the standard dataset to showcase the significance of the dataset we created. Our dataset was analyzed and compared to the IIIT-CFW dataset to demonstrate its effectiveness. The analysis revealed that the performance of all classes ranging from 10 to 50 on the UOM dataset was lower than that of the IIIT-CFW dataset. However, the UoM dataset is highly diversified and contains sketches, animated images too, making it more significant than the IIT-CFW dataset.

6 Conclusion

We have successfully integrated the deep features with the SVM classifier using the RBF kernel to recognize celebrity cartoon faces in our research paper. Our extensive experimental results have shown that SVM outperforms contemporary learning algorithms, as well as fully connected architectures like CNN and FaceNet, for recognizing celebrity cartoon faces. Subsequently, a feature selection method is employed to reduce the number of features to have an effective and efficient representation. The proposed model has been extensively tested and compared against contemporary models highlighting the superiority of our model with a minimal number of features contributing towards learning celebrity cartoon face images. Our proposed model has been extensively tested and compared against contemporary models, highlighting the superiority of our model with a minimal number of features contributing towards learning celebrity cartoon face images. We have conducted extensive experimentation to demonstrate the significance of the UoM-cartoon dataset, which consists of highly caricatured images, in comparison to the IIT-CFW dataset.

Future works

In our future work, we plan to tackle the challenge of automatically detecting and cropping cartoon faces of celebrities. Additionally, we will aim to balance the data distribution to further improve the recognition rate.

References

1. Z. Li, Comparison and Analysis of Two Cartoon Face Recognition, 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE),
2. M. Haseyama and A. Matsumura, A cartoon character retrieval system including trainable scheme, Proceedings International Conference on Image Processing (Cat. No.03CH37429),
3. Mady, Huda, and Shadi MS Hilles. Face recognition and detection using Random Forest and combination of LBP and HOG features. International conference on smart computing and electronic enterprise (ICSCEE), pp. 1-7. IEEE, (2018).
4. W. Ouarda, H. Trichili, A. M. Alimi and B. Solaiman, combined local features selection for face recognition based on Naïve Bayesian classification, 13th International Conference on Hybrid Intelligent Systems,
5. Jose, J. P., Poornima, P., & Kumar, K. M.: A novel method for color faces recognition using KNN classifier. International Conference on Computing, Communication and Applications,pp. 1-3, IEEE,(2012).
6. Maturana, Daniel, Domingo Mery, and Álvaro Soto. Face recognition with decision tree-based local binary patterns. In Asian Conference on Computer Vision, Springer, Berlin, Heidelberg, pp. 618-629, (2010).
7. Coşkun, Musab, Ayşegül Uçar, Özal Yildirim, and Yakup Demir.: Face recognition based on convolutional neural network. International Conference on Modern Electrical and Energy Systems (MEES), IEEE, pp. 376-379,(2017).
8. M. Arai, Feature extraction methods for cartoon character recognition, 5th International Congress on Image and Signal Processing,
9. X. Gao, J. Liu, J. Lin, M. Liao and L. Xiang, Contour-Preserved Retargeting of Cartoon Images, 2014 IEEE 17th International Conference on Computational Science and Engineering,
10. X. Wang and J. Yu, Learning to Cartoonize Using White-Box Cartoon Representations, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),.
11. Y. Chen, Y.-K. Lai, and Y.-J. Liu, Cartoongan: Generative adversarial networks for photo cartoonization, in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 9465–9, (2018).
12. C. Wang, J. Zhang, B. Yang, and L. Zhang, Sketch2cartoon: composing cartoon images by sketching, in Proceedings of the 19th ACM international conference on Multimedia, pp. 789–790, (2011).
13. Y. Ci, X. Ma, Z. Wang, H. Li, and Z. Luo, User-guided deep anime line art colorization with conditional adversarial networks, in Proceedings of the 26th ACM international conference on Multimedia, pp.1536–1544, (2018).
14. J. Huo, W. Li, Y. Shi, Y. Gao, and H. Yin, Web caricature: a benchmark for caricature recognition, arXiv preprint arXiv: 1703.03230, (2017).
15. A. Mishra, S. N. Rai, A. Mishra, and C. Jawahar, IIIT-CFW: A benchmark database of cartoon faces in the wild, in European Conference on Computer Vision. Springer, pp.35-47, (2016).
16. A. Fujimoto, T. Ogawa, K. Yamamoto, Y. Matsui, T. Yamasaki, and K. Aizawa, Manga109 dataset and creation of metadata, in Proceedings of the 1st international workshop on comics analysis, processing and understanding, pp.1-5,(2016).
17. M. Saito and Y. Matsui, Illustration2vec: a semantic vector representation of illustrations, in SIGGRAPH Asia 2015 Technical Briefs, pp.1-4, (2015).

18. Y. Zhou, Y. Jin, A. Luo, S. Chan, X. Xiao, and X. Yang, Toonnet: a cartoon image dataset and a cnn-based semantic classification system, in Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry, pp.1-8, (2018).
19. W. Zheng, L. Yan, F.-Y. Wang, and C. Gou, Learning from the past: Meta-continual learning with knowledge embedding for jointly sketch, cartoon, and caricature face recognition, in Proceedings of the 28th ACM International Conference on Multimedia, pp. 736-743, (2020).
20. Zheng, Y., Zhao, Y., Ren, M., Yan, H., Lu, X., Liu, J., & Li, J.: Cartoon face recognition: A benchmark dataset. In Proceedings of the 28th ACM international conference on multimedia, pp. 2264-2272,(2020).
21. Branwen, G., &Gokaslan, A.: Danbooru2019: A large-scale crowdsourced and tagged anime illustration dataset. Danbooru,(2017).
22. Jha, S., Agarwal, N., & Agarwal, S.: Bringing cartoons to life: Towards improved cartoon face detection and recognition systems. arXiv preprint arXiv:1804.01753,(2018).
23. Schroff, F., Kalenichenko, D., & Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 815-823, (2015).
24. Pal, B., Zaman, S., Hasan, A., Hasan, A. M., & Mahmud, F.: Chi-Square statistic and principal component analysis based compressed feature selection approach for Naive Bayesian Classifier. J. Artif. Intell. Res. Adv, 2, pp.16-23, (2015).
25. Mady, H., & Hilles, S. M.: Face recognition and detection using Random forest and combination of LBP and HOG features. In 2018 international conference on smart computing and electronic enterprise, ICSCEE, pp. 1-7, IEEE(2018)
26. Dino, H. I., & Abdulrazzaq, M. B.: Facial expression classification based on SVM, KNN and MLP classifiers. In 2019 International Conference on Advanced Science and Engineering (ICOASE),pp. 70-75, IEEE,(2019)
27. S. Sharma, M. Bhatt and P. Sharma.: Face Recognition System Using Machine Learning Algorithm, 5th International Conference on Communication and Electronics Systems (ICCES),
28. Jeni, L. A., Cohn, J. F., & De La Torre, F.: Facing imbalanced data--recommendations for the use of performance metrics.Humaine association conference on affective computing and intelligent interaction, pp. 245-251, IEEE (2013).
29. Liu, H., &Setiono, R.: Chi2: Feature selection and discretization of numeric attributes. In Proceedings of 7th IEEE international conference on tools with artificial intelligence, pp. 388-391, IEEE,(1995).
30. Guo, G., Li, S. Z., & Chan, K. L.: Support vector machines for face recognition. Image and Vision computing, pp.631-638,(2001).
31. Kwon, O., & Sim, J. M.: Effects of data set features on the performances of classification algorithms. Expert Systems with Applications, 40(5), pp.1847-1857, (2013).
32. Jeni, L. A., Cohn, J. F., & De La Torre, F.: Facing imbalanced data-recommendations for the use of performance metrics. Humaine association conference on affective computing and intelligent interaction pp. 245-251, IEEE (2013).