What is my name?

Nirmal Kashyap (202011031)

Supervisor : Dr. Manish K. Gupta

May 21, 2022

Content

Content

Introduction

- For a new face, we are curious about their First name, Age, Location or Ethnicity, Gender, Occupation, emotions etc.
- These are features we can extract from a facial image.
- The human brain can predict some of these attributes through a glance.
  - For instance, gender, age group, ethnicity, occupations (based on clothing in some cases), emotions (some of them), etc.
- From the features mentioned above, predicting the first name can be new. Also, nearly impossible.

Introduction

- One of the potential applications of name prediction is name association.
- Suppose, for instance, we are given a group photograph, and we know the name of the people present there.
- So using this system, one can associate the name to the face.
- Like in various social media platforms like Facebook and Instagram, people post images and tag other users.
- So, the system can associate the name to the face.

Content

Problem Statement

- Predict the first name of a person from the facial image of the person.
- The problem here is assumed to be a classification problem.
- In here, first 100 popular names of U.S. are taken.And the output will also be from the same set of these 100 names.
- The Indian dataset contains 5 names. Thus the outputs will also be from these set of 5 names.

Content

Dataset

- Name100: A First Name Face dataset[1].
- They create an extensive dataset by choosing images and tags from Flickr.
- The dataset comprises 800 faces of each 100 most famous first names, based on information from the U.S. Social Security Administration (SSA)[2].
- After completion, the dataset contains 48 men's names, 48 women-'s names, and 4 gender-neutral names.
- The names listed represent 20.35 per cent of all the Americans birthed from 1940 to 2010.

[1] Chen, Gallagher, and Girod, *"What's in a Name? First Names as Facial Attributes"*.

[2] *U.S. Social Security Administration baby name database*. URL : http://www.ssa.gov/oact/babynames.

Dataset

Figure 1: Sample of Alejandra's facial images from Name100 dataset[3].

[3] Chen, Gallagher, and Girod, *"What's in a Name? First Names as Facial Attributes"*.

Dataset

- We have created a dataset on Indian faces.
    - We tried collecting data from Google form. However, there was a lesser number of responses than we expected.
    - We have also scrapped the data from LinkedIn. The dataset comprises five classes with 180 images of each class.
    - Thus in total, this dataset has 900 images.
    - The preprocessing was done using OpenCV.
- Here we have considered five first names. There are two male names, two females names and one gender-neutral name.
- Here mainly, the first name Krishna is gender-neutral.

Dataset

Figure 2: Sample of Krishna's facial images from Indian dataset.

Figure 3: Sample of Pranav's facial images from Indian dataset.

Content

Experiments and Results

- We have implemented various algorithms. We have implemented Vision Transformer [8], ensemble learning like Bagging [1] , and boosting algorithms like Adaboost [6] and XGBoost [4].
- For extracting features, we have used Spatial pyramid pooling [7] and Deep Learning features. For Deep learning, it uses the Relu activation function.
- Vision Transformer
  - Vision transformer [8] is the state of the art for image classification for ImageNet with given conditions.
  - We have used ViT_B_16 with the image resolution of 224 x 224, and it was pretrained on the Imagenet21k dataset. Here B stands for Base. 16 is the patch size.

Experiments and Results

- The ensemble learning technique combines various weak learners to achieve higher prediction accuracy.
- These weak learners can be any classifier algorithm. Some of the commonly used learners are Random forests.
- A few of the commonly used Ensemble learning techniques are bagging and boosting.
    - In the bagging technique [1], all the classifiers learn independently and parallelly. The results of all these classifiers are combined to determine the final output.
    - In boosting technique, classifiers are used in an iterative manner. Here it learns from its previous mistakes. XGBoost [8] and AdaBoost [6] are boosting algorithms.

Table 1: The numbers represent the accuracy. The models where the accuracy numbers are higher than the baseline have been represented in bold.

Sr. No.	Model	Name100 Subset 1	Name100 Subset 2	Indian
1	ViT [8]	35%	44.2%	41.67%
2	Bagging SPP	35.63%	37.38%	42.2%
3	Adaboost SPP	34.63%	37.38%	41.6%
4	XGBoost SPP	32.75%	37.25	44.44%
5	Adaboost DL	27.5%	31.62%	39.44%
6	Bagging DL	29.67%	31.62%	39.44%

In the table 1 prefix SPP stands for Spatial pyramid pooling. These algorithms have used SPP for feature extraction. While the suffix DL stands for Deep learning feature.

Experiments and Results

- MFSVM [3] is treated as the baseline. It uses boosting technique AdaBoost [6] with the classifier SVM.
- The baseline for five classes is 39.4%. With ViT [8], we achieved higher accuracy than baseline for certain subsets of the dataset.
- While for 10 classes, the baseline is 23.5%. With ViT [8] for certain subsets of the dataset, it achieves an accuracy of 26%.
- Name100 subset 1 stands for classes Chris, Jamie, Maggie, Stephen, and Tina.
- Name100 subset 2 stands for classes Abby, Amanda, Angela, Aaron, and Andrea.

Experiments and Results

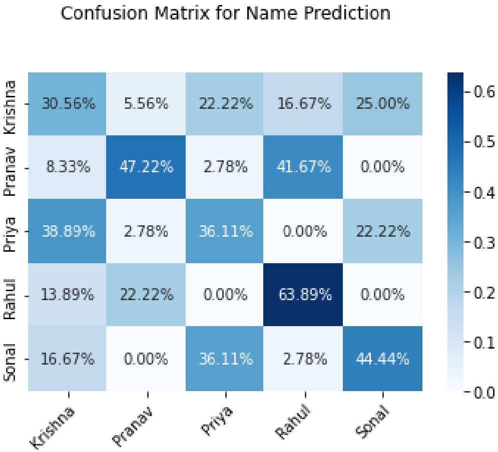

Figure 4: Confusion matrix for Indian Dataset on XGBoost SPP model. The overall accuracy is 44.44% here.

Experiments and Results

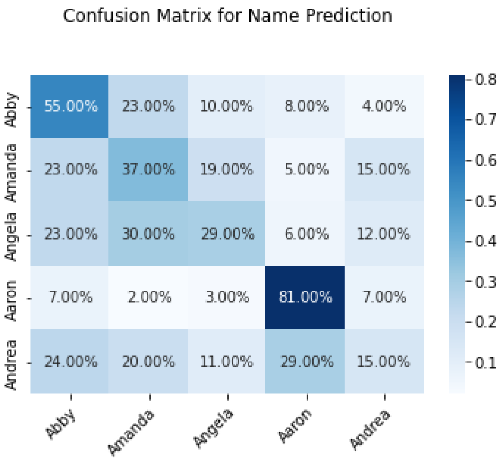

Figure 5: Confusion matrix for Name100 Dataset for 5 classes Subset 2. The overall accuracy is 44.12% here.

Experiments and Results

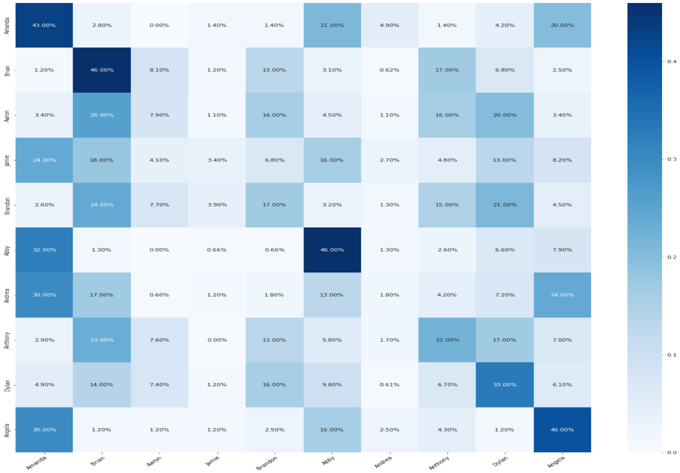

Figure 6: Confusion matrix for Name100 Dataset for 10 classes. The overall accuracy is 26%.

Experiments and Results

- From the results of our experiments we found that the first name is mainly misclassified among the name of same-gender itself. [4]
- Male names are majorly misclassified into male names only. Rarely for some classes and some models, there can be some exceptions.
- For the gender-neutral names like Krishna and Jamie, it is non-deterministic.
- The overall accuracy for almost all the classes is majorly higher than the random chances with some exceptions.

[4] *The resuling confusion matrix of all the other models from table 1 can be seen here. Link.* https://docs.google.com/presentation/d/1ca1Bc9wPzgsABDhRIupe-HTFolChkBE4L7EoT1SXNkE/edit?usp=sharing

Content

Conclusion

- The name prediction system is not that ideal to function well in all real-world situations.
- We have constructed the dataset for Indian faces. And also concluded that name prediction of Indian faces is also possible.
- The current works have high prediction accuracy than random chance.
- The classes with gender-neutral names would require certain amount of future efforts to increase the prediction accuracy.
- But still plenty of work needs to be carried out to achieve higher efficiency and accuracy goals.

Content

Future Work

- We would want to expand the size of the custom Indian dataset.
  - We want to increase the number of classes, i.e. various unique first names.
  - Moreover, we would like to increase the number of image samples per class.
- We also would like to see and verify if this pattern exists for datasets of other nationalities.
- We would also like to improve the overall classification accuracy.
- We would specifically like to increase the prediction accuracy for gender-neutral classes.
- We would also like to work in some application domain of the name prediction.

References I

[1] Leo Breiman. "Bagging predictors". In: *Machine learning* 24.2 (1996), pp. 123–140.

[2] Angus Chen. *Your Name Might Shape Your Face, Researchers Say*. 2017. URL: https://www.npr.org/sections/health-shots/2017/02/27/517496915/your-name-might-shape-your-face-researchers-say.

[3] Huizhong Chen, Andrew Gallagher, and Bernd Girod. "What's in a Name? First Names as Facial Attributes". In: *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (June 2013), pp. 3366–3373. DOI: 10.1109/CVPR.2013.432.

[4] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[5] *Faceprint definition & meaning*. URL: https://www.dictionary.com/browse/faceprint.

[6] Trevor Hastie et al. "Multi-class adaboost". In: *Statistics and its Interface* 2.3 (2009), pp. 349–360.

[7] Kaiming He et al. "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition". In: *CoRR* abs/1406.4729 (2014). arXiv: 1406.4729. URL: http://arxiv.org/abs/1406.4729.

References III

[8] Alexander Kolesnikov et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". In: 2021.

[9] Robin Kramer and Alex L. Jones. *Do people's first names match their faces?* 2015. URL: https://www.jasnh.com/pdf/Vol12-No1-article1.pdf.

[10] *U.S. Social Security Administration baby name database*. URL: http://www.ssa.gov/oact/babynames.

Thank You

MFSVM pseudo code

- Data: Training data $x_{t,i}$, training labels $y_i \in \{-1, +1\}$,testing data z_t, where $t = 1, \ldots, T$ and $i = 1, \ldots, N$
- Result: SVM classifiers $f_t(z_t)$, classifier weights α_t
- Initialization: weights $D_i = 1$
- for $t = 1 : T$ do

  (i) Do SVM 5-fold cross validation with weights D for obtaining confidence $f_t^{cv}(x_{t,i}) \in R$ and prediction $\hat{y}_{t,i}^{cv} = \text{sign}(f_t^{cv}(x_{t,i}))$, calculate error $err_t = \frac{\sum_{t=1}^{N} \left| \left\{ \hat{y} | \hat{y}_{t,i} \neq y_i \right\} \right|}{N}$

  (ii) Train SVM f_t with D;

  (iii) Calculate $\alpha_t = \frac{1}{2} \log\left(\frac{1-err_t}{err_t}\right)$;

  (iv) Set $D_i = D_i \exp(-\alpha_t y_i f_t^{cv}(x_{t,i}))$, and renormalize so that $\sum_{t=1}^{N} D_i = N$;

  end
- Output the final classifier $f_{\text{all}}(z) = \sum_{t=1}^{N} \alpha_t f_t(z_t)$

Algorithm: Multi-Feature SVM[2]