B.TECH PROJECT REPORT-II

Yoga Guru: A Study of Pose Estimation and Machine and Deep Learning Methods For Yoga Pose Correction

Submitted by:
Dhananjay Sharma (2K18/CO/119)
Harshil Panwar (2K18/CO/142)
Harshit Goel (2K18/CO/143)

Under Guidance of:
Prof. Rahul Katarya
(Supervisor)

INDEX

Proof of Conference

GL BAJAJ
Institute of Technology & Management
FIND YOUR SPARK
Approved by AICTE & Affiliated to AKTU

INTERNATIONAL CONFERENCE
on
INTELLIGENT COMPUTING AND SECURITY: A PARADIGM SHIFT
(IICS2021)
Organised by
Department of Master of Computer Application
G.L. Bajaj Institute of Technology & Management, Greater Noida
December 17th - 18th, 2021

Springer
Scopus

CALL FOR PAPERS | IMPORTANT DATES | PAPER SUBMISSION | REGISTRATION | SPEAKERS | CONTACT US

- HOME
- ABOUT THE COLLEGE
- PATRONS & CHAIRS
- EDITORIAL BOARD
- ADVISORY COMMITTEE
- ORGANIZING COMMITTEE
- TECHNICAL PROGRAM COMMITTEE

Important Dates

Paper Submission Started:	15 September, 2021
Submission deadline:	15 November, 2021
Notification of acceptance:	5 December, 2021
Submission deadline for revised papers:	10 December, 2021
Date of Conference:	17-18 December, 2021

GL BAJAJ
Institute of Technology & Management
FIND YOUR SPARK
Approved by AICTE & Affiliated to AKTU

Springer

INTERNATIONAL CONFERENCE
on
INTELLIGENT COMPUTING AND SECURITY:
A PARADIGM SHIFT (IICS2021)
December 17th - 18th 2021

CERTIFICATE

Certified that Ms./Mr./Dr. Dhananjay Sharma

from Delhi Technological University, Delhi

has participation / presented paper entitled Deep-Learning for Self-Learning in Yoga and Fitness: A Literature Review

in Two Days Springer Sponsored Intelligent Computing and Security: A Paradigm Shift (IICS2021) on December 17th - 18th 2021 *organized by Department of Master of Computer Applications, G.L. Bajaj Institute of Technology & Management Greater Noida, (U.P.) India.*

Convener	Conference Chair	General Chair
Dr. Sanjeev Kumar	Prof. (Dr.) Madhu Sharma Gaur	Prof (Dr.) Rajeev Agrawal

Problem Statement

01 HEALTH

With a rise in stress levels, obesity rates and health issues across the globe it has become imperative that we incorporate exercising as a part our daily routine

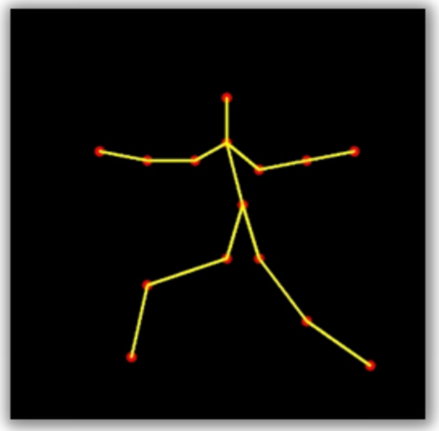
02 SCHEDULE

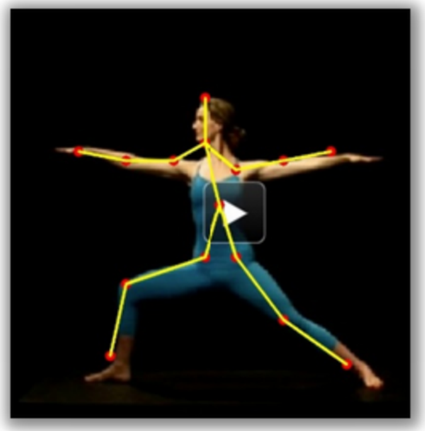
It is extremely difficult to afford a gym membership/private coach and match our schedule according to theirs on a daily basis.

03 PROPOSAL

Thus, we decided to propose a Deep Learning based solution that could solve the above problem.
A pipeline capable of analyzing a person's exercise posture and returning a feedback score!

Generic Pipeline





Dataset

- Five Yoga Pose Output Classes,
- Position and Color Augmentation to increase training data.
- Total 2000 images manually scraped
- Experimentation with 3 different dataset classes.

Pose Estimation

01 Backbone Architectures

Range from AlexNet, Fast R-CNN, Mask R-CNN, to feature pyramid networks. VGG and ResNet remain popular choice.

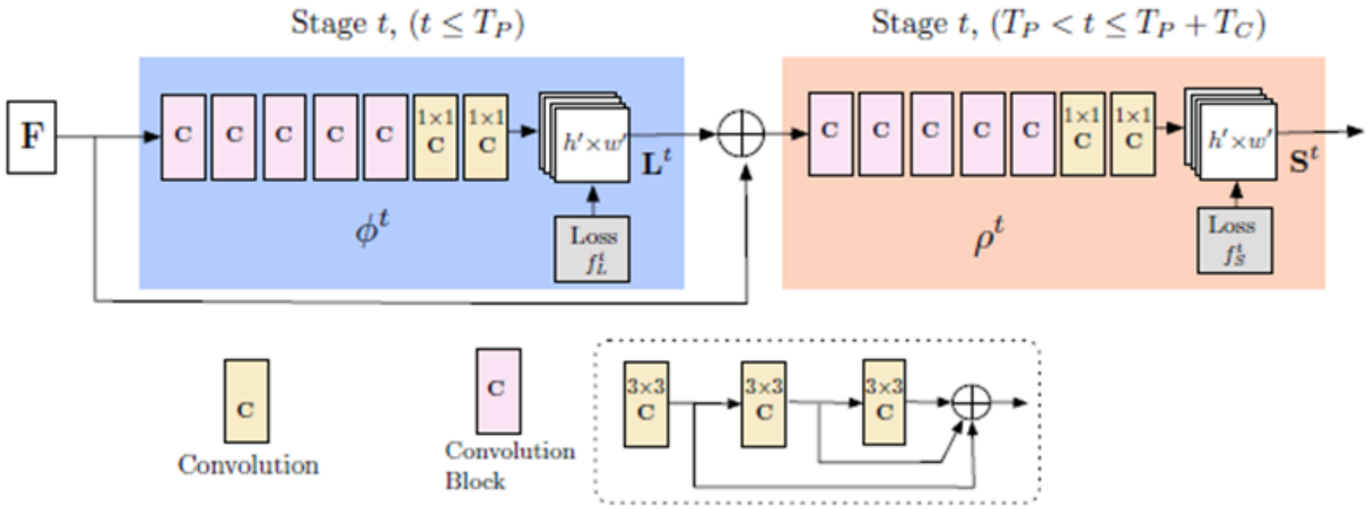
02 Loss Functions

Cross-Entropy loss, Mean Absolute Error, and Mean Squared Error are mostly used.

03 Evaluation Metrics

- Percentage of Correct Parts
- Percentage of Detected Joints
- Percentage of Correct Key-Points

Model	Year	Backbone Architecture	Approach	Loss Function	Evaluation Metrics	Training Datasets
DeepPose [30]	2014	AlexNet	Top-down	L2	PDJ, PCP	LSP, FLIC
DeeperCut [31]	2016	ResNet	Bottom-Up	L1, Cross-Entropy	AP, mAP, AUC, PCKh@0.5	MPII, COCO, LSP
Convolutional Pose Machines	2016	VGG	Top-down	L2	PCKh@0.1,PCKh@0.2, PCKh@0.5	FLIC, LSP, MPII
Rmpe: Regional Multi-Person Pose Estimation [32]	2017	VGG, ResNet	Top-down	L2	mAP	MPII, COCO
DensePose [33]	2018	FCN, Mask-RCNN	Bottom-up	Cross-Entropy	AP, PCP, GPS	COCO
High-Resolution Network (HRNet) [34]	2019	ResNet	Bottom-up	L2	AP, mAP, PCKh@0.5	MPII, COCO
Human Pose Estimation for Real-World Crowded Scenarios [35]	2019	ResNet	Top-down	L2	AP, OKS	CrowdPose, JTA

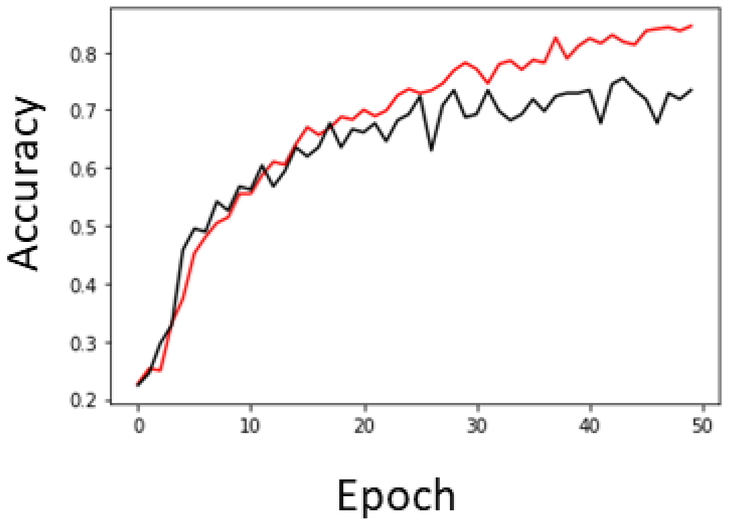


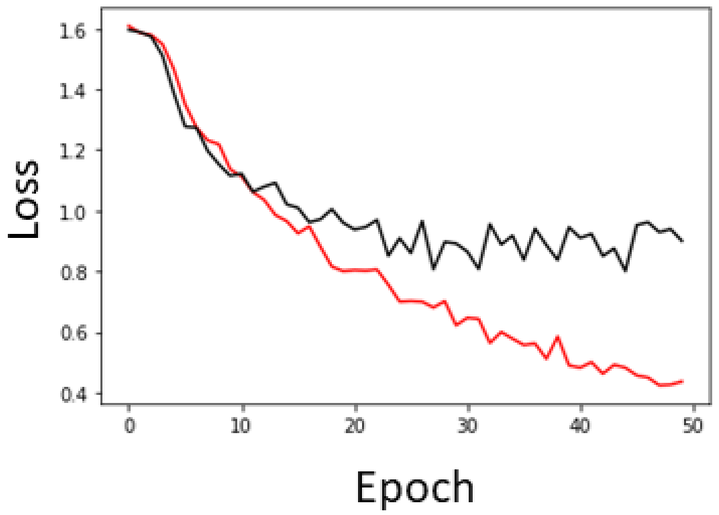
OpenPose Multi-Stage CNN Architecture

Estimating joint angles based on OpenPose indexing

Angle Name	Body Parts			MPII Label Index
elbow-r	Right Shoulder	Right Elbow	Right Wrist	[2,3,4]
elbow-l	Left Shoulder	Left Elbow	Left Wrist	[5,6,7]
knee-r	Right Hip	Right Knee	Right Ankle	[8,9,10]
knee-l	Left Hip	Left Knee	Left Ankle	[11,12,13]
neck-r	Head	Neck	Right Shoulder	[0,1,2]
neck-l	Head	Neck	Left Shoulder	[0,1,5]
leg-r	Chest	Right Hip	Right Knee	[14,8,9]
leg-l	Chest	Left Hip	Left Knee	[14,11,12]
back	Head	Neck	Chest	[0,1,14]

CLASSIFICATION TECHNIQUES

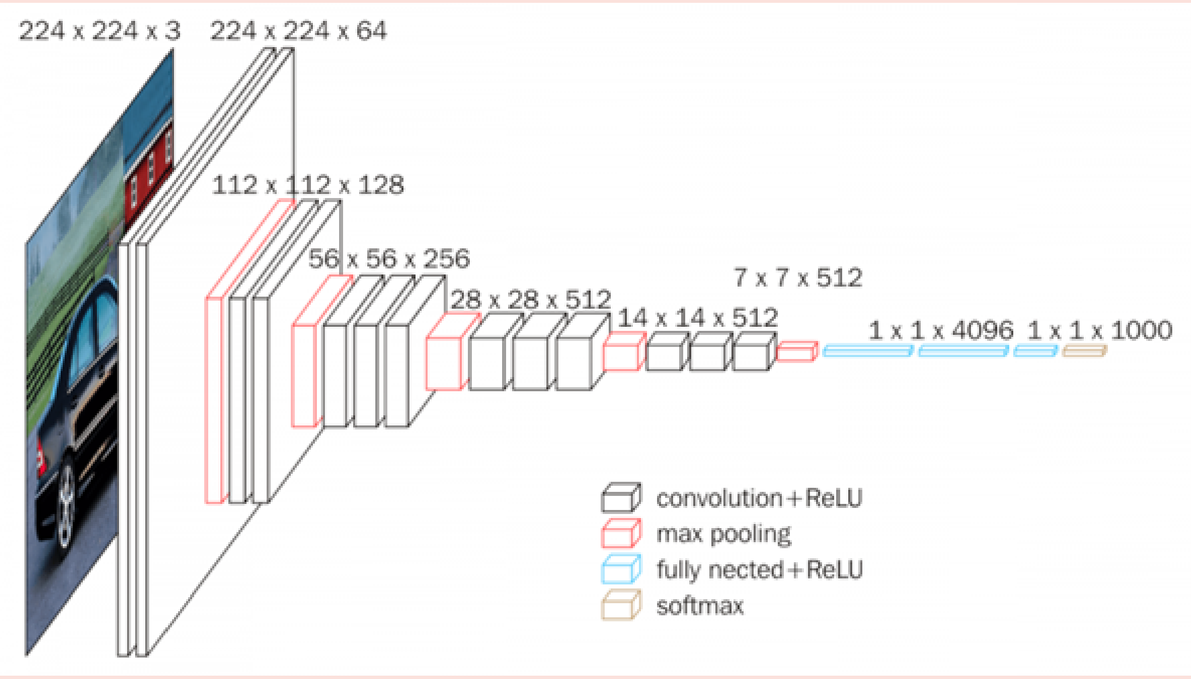


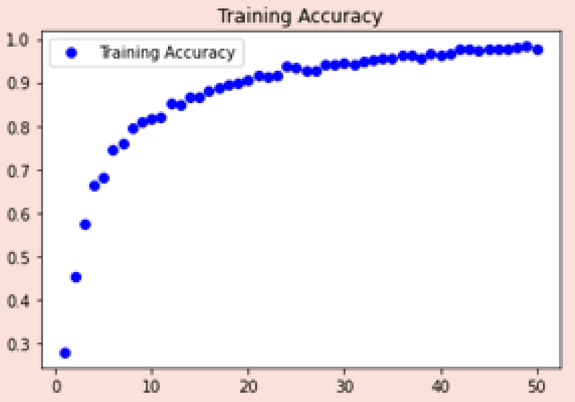


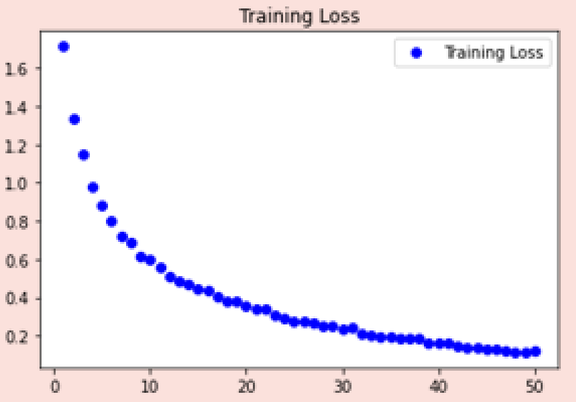

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 148, 148, 32)      896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 74, 74, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 72, 72, 64)        18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 36, 36, 64)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 34, 34, 128)       73856
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 17, 17, 128)       0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 15, 15, 128)       147584
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 7, 7, 128)         0
_________________________________________________________________
flatten (Flatten)            (None, 6272)              0
_________________________________________________________________
dense (Dense)                (None, 64)                401472
_________________________________________________________________
dropout (Dropout)            (None, 64)                0
_________________________________________________________________
dense_1 (Dense)              (None, 5)                 325
=================================================================
Total params: 642,629
Trainable params: 642,629
Non-trainable params: 0
_________________________________________________________________
```

Custom Convolutional Neural Network

TRANSFER LEARNING: VGG - 16



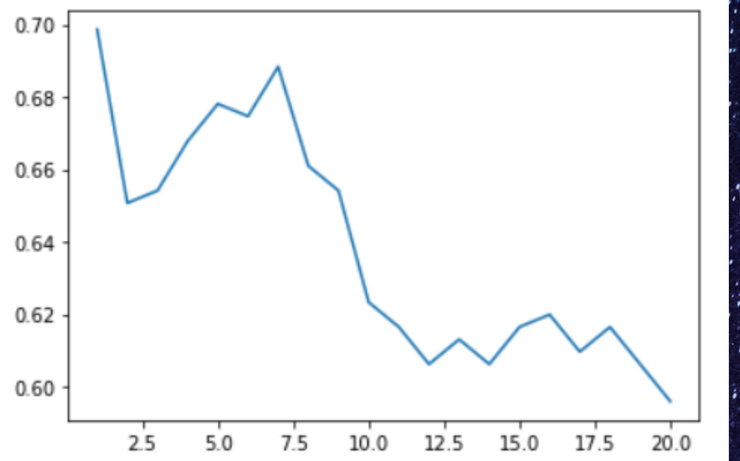




OTHER CLASSIFIERS

```
sns.lineplot(x = nb, y = acc)
```

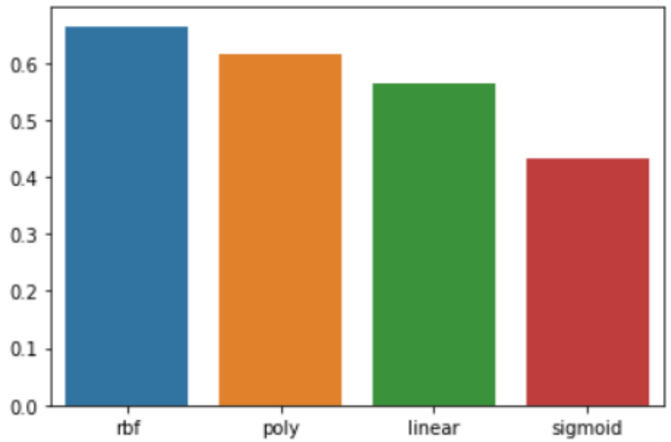
```
<AxesSubplot:>
```



KNN (Acc v/s N)

```
sns.barplot(x = kernels, y=acc)
```

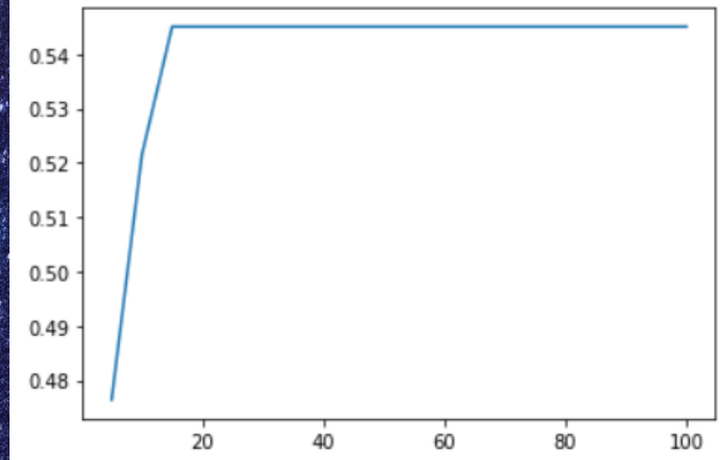
```
<AxesSubplot:>
```



SVM (Acc v/s Kernel)

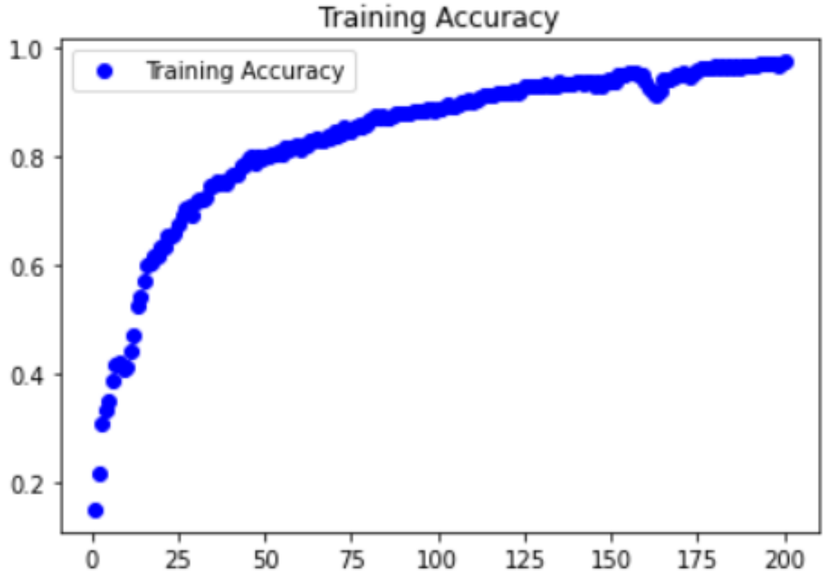
```
sns.lineplot(x = md, y = acc)
```

```
<AxesSubplot:>
```



DT (Acc v/s Depth)

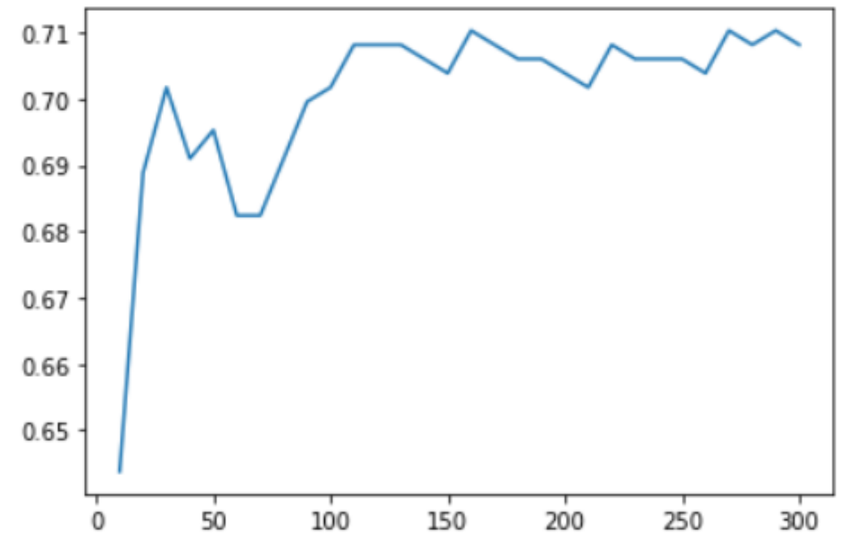
```
plt.show()
```



ANN (Acc v/s Epochs)

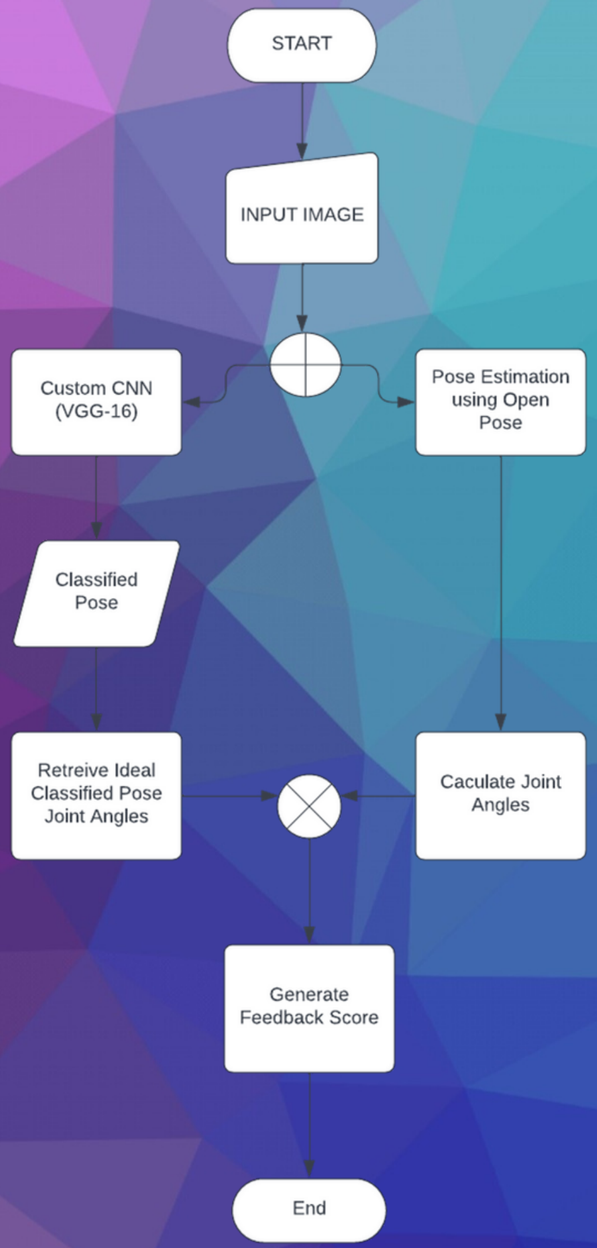
```
sns.lineplot(x = ne, y = acc3)
```

```
<AxesSubplot:>
```

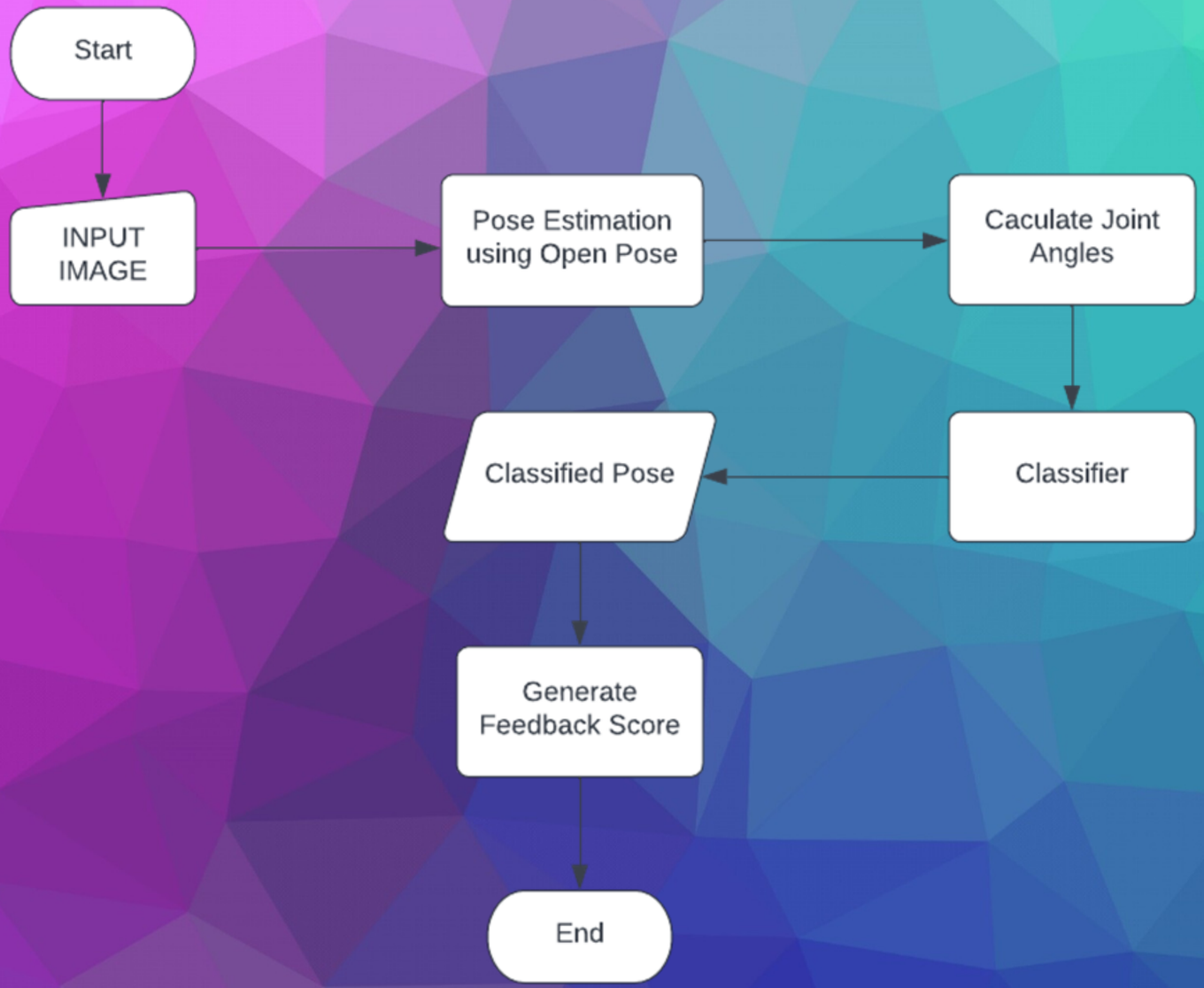


RF (Acc v/s N Trees)

THE TWO FLOWS



YOGA GURU FULL SCALE PIPELINE



YOGA GURU LIGHTWEIGHT PIPELINE

RESULTS

BY THE NUMBERS

Datasets	Train Accuracy	Test Accuracy
Plain Images	98.32%	93.56%
Pose Estimation on black background	85.21%	56.44%
Super imposed Images	98.06%	91.47%

Dataset Wise CNN

Model Structure	Train Accuracy	Test Accuracy
ReLu(16,12), Softmax(5)	73%	64%
ReLu(16,12, 8), Softmax(5)	78%	64%
ReLu(20, 16,12, 8), Softmax(5)	85%	64%
ReLu(20, 20, 16,12, 8), Softmax(5)	92%	66%
ReLu(24, 20, 20, 16,12, 8), Softmax(5)	97%	65%

Various ANNs

Algorithm	Best Test Accuracy
KNN	69.86%
SVM	66.43%
ANN	66.78%
Decision Tree	54.50%
Random Forest	71.03%

Best Test Accuracies

YOGA GURU




Here is Your Output

Predicted Pose	Warrior2

Score	8	10

Try Again ?

Output in the Flask App

Conclusion

1. Highlighted the importance and need for such a product
2. Summarised cutting edge pipelines, pose detection algorithms
3. Experimented with and analyzed various Machine learning and deep learning algorithms to suit our needs
4. Presented two product alternatives according to user's needs and equipments

References

1.T. L. Munea, Y. Z. Jembre, H. T. Weldegebriel, L. Chen, C. Huang and C. Yang, "The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation," in IEEE Access, vol. 8, pp. 133330-133348, 2020, doi: 10.1109/ACCESS.2020.3010248.

2. Cao, Zhe & Hidalgo, Gines & Simon, Tomas & Wei, Shih-En & Sheikh, Yaser. (2018). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.

3. Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018).

4. Hochreiter, Sepp & Schmidhuber, Jürgen. (1997). Long Short-term Memory. Neural computation. 9. 1735-80. 10.1162/neco.1997.9.8.1735.

5. Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems. Curran Associates, Inc..

6. Girshick, Ross. "Fast r-cnn." Proceedings of the IEEE international conference on computer vision. 2015.

7. He, Kaiming & Gkioxari, Georgia & Dollár, Piotr & Girshick, Ross. (2017). Mask R-CNN

8. Lin, Tsung-Yi, et al. "Feature pyramid networks for object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

9. VGG - Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

10. ResNet - He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

References

11. J. Palanimeera, K. Ponmozhi, Classification of yoga pose using machine learning techniques, Materials Today: Proceedings, Volume 37, Part 2, 2021, Pages 2930-2933, ISSN 2214-7853

12. Guo, Gongde & Wang, Hui & Bell, David & Bi, Yaxin. (2004). KNN Model-Based Approach in Classification.

13. Evgeniou, Theodoros & Pontil, Massimiliano. (2001). Support Vector Machines: Theory and Applications. 2049. 249-257. 10.1007/3-540-44673-7_12.

14. Peng, Joanne & Lee, Kuk & Ingersoll, Gary. (2002). An Introduction to Logistic Regression Analysis and Reporting. Journal of Educational Research - J EDUC RES. 96. 3-14. 10.1080/00220670209598786.

15. Rish, Irina. (2001). An Empirical Study of the Naïve Bayes Classifier. IJCAI 2001 Work Empir Methods Artif Intell. 3.

16. M. Gochoo et al., "Novel IoT-Based Privacy-Preserving Yoga Posture Recognition System Using Low-Resolution Infrared Sensors and Deep Learning," in IEEE Internet of Things Journal, vol. 6, no. 4, pp. 7192-7200, Aug. 2019, doi: 10.1109/JIOT.2019.2915095.

17. M. C. Thar, K. Z. N. Winn and N. Funabiki, "A Proposal of Yoga Pose Assessment Method Using Pose Detection for Self-Learning," 2019 International Conference on Advanced Information Technologies (ICAIT), 2019, pp. 137-142, doi: 10.1109/AITC.2019.8920892.

18. F. Rishan, B. De Silva, S. Alawathugoda, S. Nijabdeen, L. Rupasinghe and C. Liyanapathirana, "Infinity Yoga Tutor: Yoga Posture Detection and Correction System," 2020 5th International Conference on Information Technology Research (ICITR), 2020, pp. 1-6, doi: 10.1109/ICITR51448.2020.9310832.

19. M. Verma, S. Kumawat, Y. Nakashima and S. Raman, "Yoga-82: A New Dataset for Finegrained Classification of Human Poses," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, pp. 4472-4479, doi: 10.1109/CVPRW50498.2020.00527.

References

20. R. Huang, J. Wang, H. Lou, H. Lu and B. Wang, "Miss Yoga: A Yoga Assistant Mobile Application Based on Keypoint Detection," 2020 Digital Image Computing: Techniques and Applications (DICTA), 2020, pp. 1-3, doi: 10.1109/DICTA51227.2020.9363384.

21. I. Gregory and S. M. Tedjojuwono, "Implementation of Computer Vision in Detecting Human Poses," 2020 International Conference on Information Management and Technology (ICIMTech), 2020, pp. 271-276, doi: 10.1109/ICIMTech50083.2020.9211145.

22. E. W. Trejo and P. Yuan, "Recognition of Yoga Poses Through an Interactive System with Kinect Device," 2018 2nd International Conference on Robotics and Automation Sciences (ICRAS), 2018, pp. 1-5, doi: 10.1109/ICRAS.2018.8443267.

23. G. Dsouza, D. Maurya and A. Patel, "Smart gym trainer using Human pose estimation," 2020 IEEE International Conference for Innovation in Technology (INOCON), 2020, pp. 1-4, doi: 10.1109/INOCON50539.2020.9298212.

24. Y. Hou, H. Yao, H. Li and X. Sun, "Dancing like a superstar: Action guidance based on pose estimation and conditional pose alignment," 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 1312-1316, doi: 10.1109/ICIP.2017.8296494.

25. Wei, Shih-En & Ramakrishna, Varun & Kanade, Takeo & Sheikh, Yaser. (2016). Convolutional Pose Machines.

26. Ren, Shaoqing & He, Kaiming & Girshick, Ross & Sun, Jian. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 39. 10.1109/TPAMI.2016.2577031.

References

27. A. Nagarkoti, R. Teotia, A. K. Mahale and P. K. Das, "Realtime Indoor Workout Analysis Using Machine Learning & Computer Vision," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 1440-1443, doi: 10.1109/EMBC.2019.8856547.

28. Jianbo Wang, Kai Oiu Houwen Peng, Jianlong Fu, and Jianke Zhu. 2019. AI Coach: Deep Human Pose Estimation and Analysis for Personalized Athletic Training Assistance. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19).

29. Bertinetto, Luca & Valmadre, Jack & Henriques, Jao & Vedaldi, Andrea & Torr, Philip. (2016). Fully-Convolutional Siamese Networks for Object Tracking. 9914. 850-865. 10.1007/978-3-319-48881-3_56.

30. A. Toshey and C. Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks," 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp, 1653-1660, doi: 10.1109/CVPR.2014.214.

31. Insafutdinov, Eldar & Pishchulin, Leonid & Andres, Bioern & Andriluka, Mykhaylo & Schiele, Bernt. (2016). DeeperCut: A Deeper, Stronger, and Faster Multi-person Pose Estimation Model. 9910. 34-50. 10.1007/978-3-319-46466-4_3.

32. Fang, Hao-Shu & Xie, Shuqin & Tai, Yu-Wing & Lu, Cewu (2017). RMPE: Regional Multi-person Pose Estimation. 2353-2362. 10.1109/ICCV.2017.256.

33. R. A. Güler, N. Neverova and I. Kokkinos, "DensePose: Dense Human Pose Estimation in the Wild," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7297-7306, doi: 10.1109/CVPR.2018.00762.

34. Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W., & Xiao, B. (2020). Deep High-Resolution Representation Learning for Visual Recognition.

References

35. T. Golda, T. Kalb, A. Schumann and J. Beyerer, "Human Pose Estimation for Real-World Crowded Scenarios," 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2019, pp. 1-8, doi: 10.1109/AVSS.2019.8909823.

THANK YOU !

Submitted by:
Dhananjay Sharma (2K18/CO/119)
Harshil Panwar (2K18/CO/142)
Harshit Goel (2K18/CO/143)