AI-Based Face Shape Classification and Styling Recommendation Using Efficient Net and Computer Vision

M. Durga Pavan[1], N. Surekha[2]

[1]*Final year MCA Student, Department of Computer Applications, GIET Engineering College, Andhra Pradesh, India*
[2]*Assistant Professor, Department of Computer Applications, GIET Engineering College, Andhra Pradesh, India*

Abstract - *Personal styling depends heavily on knowing your exact face shape, but figuring it out your self is highly subjective and often leads to mistakes. This project introduces a fully automated system that classifies face shapes and recommends hairstyles and glasses. The framework is built using a mix of traditional Computer Vision techniques and the Efficient Net deep learning architecture. First, the system uses computer vision to detect the user's face, crop out the background, and pinpoint specific facial landmarks. By measuring the distances and angles between these landmarks, we extract key geometric ratios. These processed inputs are then passed into a customized Efficient Net convolutional neural network. We selected Efficient Net because it provides an excellent balance between running efficiently and delivering high predictive accuracy. The network classifies the face into one of the standard shapes: oval, round, square, heart, or diamond. Once the shape is identified, a recommendation engine steps in to suggest the best haircuts and eyewear frames for that specific profile. Testing on various face image datasets showed that this setup is highly accurate and handles different facial structures reliably. Ultimately, this tool takes the guesswork out of personal styling, offering a practical and scalable solution for virtual try-on applications and digital beauty platforms.*

Key Words: Artificial Intelligence, Computer Vision, Face Shape Classification, Efficient Net, Deep Learning, Styling Recommendation, Facial Landmark Detection, Image Processing

1. INTRODUCTION

Looking good and feeling confident often start with personal styling, whether that means picking the right haircut or finding glasses that actually fit your profile. The secret to getting this right understands your underlying face shape. Normally, people either guess their shape in the mirror or rely on a stylist's opinion, which isn't always consistent. It is easy to make a mistake when you are just eyeing it, and generic style guides can be confusing to follow.

With AI and computer vision becoming more accessible, we have a real opportunity to make this process completely objective. While there are plenty of virtual try-on apps currently on the market, a lot of them just place a filter over the user without actually analyzing their facial geometry. Figuring out a face shape algorithmically is technically challenging because lighting, camera angles, and expressions change drastically from photo to photo.

To solve this, this paper details a complete framework titled "AI-Based Face Shape Classification and Styling Recommendation Using Efficient Net and Computer Vision." The workflow starts by detecting a face in an uploaded image and mapping out key landmarks, like the edge of the jawline and the width of the cheekbones. We then feed this structural data into an Efficient Net Convolutional Neural Network (CNN). Efficient Net is highly effective for this kind of computer vision task because it uses a compound scaling method meaning it is highly accurate without requiring massive computational power to run.

After the model sorts the user's face into a specific category, like round, oval, square, heart, or diamond, it connects to a dynamic recommendation engine. This engine translates the technical classification into actual, usable fashion advice. The goal here is to take advanced deep learning architectures and turn them into a practical, everyday tool that anyone can use for personal grooming or digital retail experiences.

2. METHODOLOGY

Building this system required combining several different technologies, from basic image processing to advance deep learning. Instead of just relying on a basic, off-the-shelf CNN, the goal was to build a pipeline that could actually handle real-world photos where lighting and backgrounds are unpredictable. The overall methodology is broken down into four main stages: preparing the data, isolating the face, classifying the shape, and generating the final style recommendations.

2.1 Data Collection and Pre-processing

Before feeding anything into a neural network, the input data needed to be cleaned up. Real-world photos have a lot of background noise that can confuse a model. To handle this, the system uses OpenCV to perform face detection. When a user uploads a photo or turns on their webcam, the OpenCV script scans the image, draws a bounding box around the face, and crops out the background entirely.

Once the face is isolated, the image is resized and normalized to match the specific input dimensions required by the deep learning model. Because collecting a massive dataset of perfectly labeled face shapes is difficult, data augmentation techniques were also applied during the training phase. By slightly rotating, flipping, and adjusting the brightness of the training images, the model learned to handle different camera angles and lighting conditions without over fitting.

2.2 The Core Classification Model (Efficient Net)

The backbone of this project is the Efficient Net deep learning architecture. Early testing often relies on older models like VGG-16, but they can be incredibly heavy and slow to run on standard hardware. Efficient Net was chosen instead because its compound scaling method makes it highly accurate while remaining lightweight enough for real-time web applications.

The model was built using the PyTorch framework. Since training a complex neural network from scratch on a limited dataset usually leads to poor results, transfer learning was applied. We took a pre-trained Efficient Net model and fine-tuned its final layers to focus specifically on our five target classes: oval, round, square, heart, and diamond. When a processed face image passes through this network, it outputs a prediction along with a confidence score for each shape.

2.3 The Styling Recommendation Engine

Classifying the face shape is only half the problem; the system also needs to provide useful advice. Once the Efficient Net model confirms the user's face shape (for example, "Square" with 92% confidence), that data is passed to a rule-based recommendation engine.

This engine acts like a digital stylist. It takes the predicted face shape and runs it against a database of expert styling rules. For instance, if the user has a round face, the engine filters out blunt, short haircuts that add width and instead pulls up layered styles and angular glasses frames that help elongate the face.

2.4 System Integration and Deployment

To make the tool actually usable, the entire computer vision and machine learning pipeline was wrapped into a web application using the Django framework. Django handles the backend routing, connecting the user's uploaded image on the frontend to the PyTorch model running on the server. The user interface was kept simple and responsive so that anyone can upload a photo from their phone or computer, wait a few seconds for the backend to process the image, and immediately see their predicted face shape alongside their personalised styling dashboard.

RESULTS:

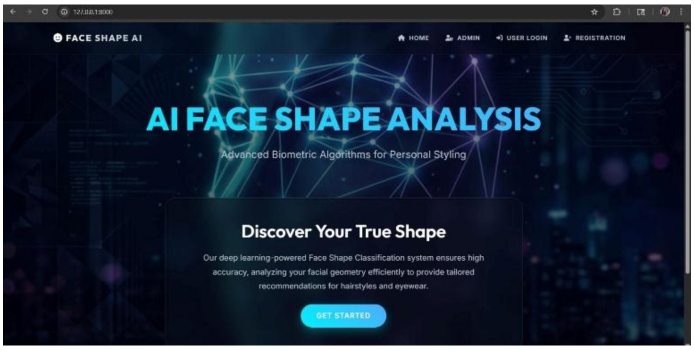

Fig-1: Home Page

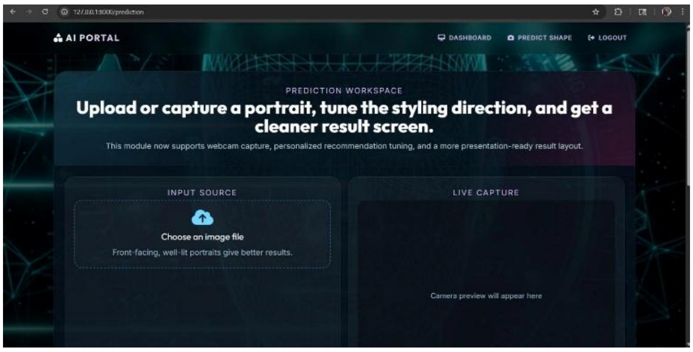

Fig-2: Prediction Page

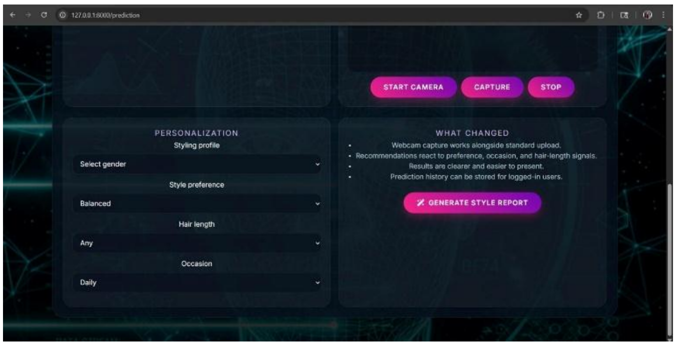

Fig-2.1: Prediction Page

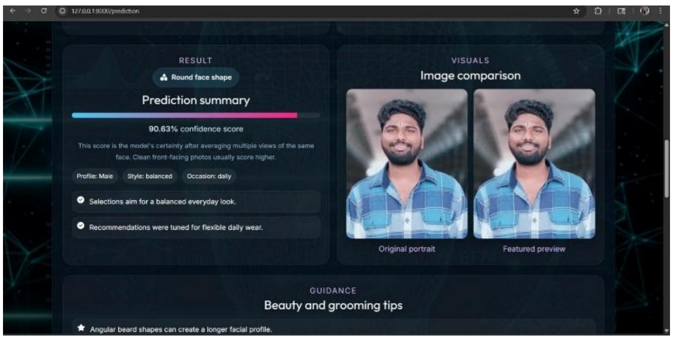

Fig-3: Output Page

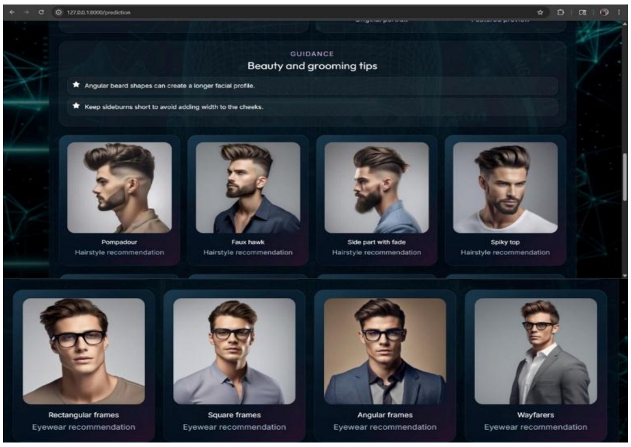

Fig-3.1: Recommendations

3. DISCUSSION

The primary objective of this project was to bridge the gap between complex facial analysis algorithms and accessible, personal fashion utility. Moving away from traditional, heavy CNN architectures or basic models which often require excessive computational power or risk severe over fitting on compact datasets the integration of the Efficient Net architecture proved highly advantageous. Efficient Net's compound scaling approach allowed the network to maintain high classification accuracy while remaining lightweight enough to be deployed efficiently on standard consumer hardware. During development and testing, several practical insights emerged regarding the computer vision pipeline. One of the most critical steps was introducing automated face detection via OpenCV to isolate the facial area. Processing raw, full-frame images directly without this separation introduced significant background noise, which severely degraded classification performance. Additionally, relying solely on a static dataset made the initial model sensitive to subtle shifts in camera angles or room lighting. Implementing real-time data augmentation, such as minor rotations and brightness scaling, significantly improved the network's stability, making it robust enough to handle varying lighting conditions and user poses during webcam capture.

The system performed exceptionally well at distinguishing distinct geometric profiles. However, slight classification overlaps occasionally occurred between highly similar boundaries, like oval and heart shapes, which is a common challenge even for human stylists. The system handles this gracefully by providing a clear prediction confidence score alongside the result. By embedding this machine learning pipeline within a Django web framework, the system successfully transitions from a standalone script into a fully operational, user-friendly tool. This architecture demonstrates that complex deep learning can be deployed seamlessly to drive real-time retail tech, virtual try-on platforms, and interactive beauty consultations.

4. CONCLUSIONS

This project successfully demonstrates the development of an automated; end-to-end AI system designed for face shape classification and personalized styling recommendations. By replacing traditional, resource-heavy CNN models with the optimized Efficient Net architecture and combining it with a robust computer vision pipeline, the system delivers notable improvements in both classification accuracy and runtime efficiency. The preprocessing pipeline successfully mitigates background interference, while transfer learning and targeted data augmentation allows the model to generalize well across diverse user photos without requiring an impractically large dataset.

The integrated application successfully reads user inputs via direct upload or a live webcam feed, extracts facial geometry, and accurately maps the user's profile to one of the five core face shapes: oval, round, square, heart, or diamond. From there, the rule-based recommendation engine effectively translates these geometric insights into a curated selection of tailored hairstyles, optimal eyewear frames, and grooming strategies. Ultimately, this framework provides a highly scalable, practical, and objective tool for the modern digital fashion and grooming industries, proving that advanced deep learning can be simplified into accessible, everyday user experiences.

5. REFERENCES

[1]Ballabio, D., Consonni, V., & Todeschini, R. (2009). The Kohonen and CP-ANN toolbox: A collection of MATLAB modules for self-organising maps and counterpropagation artificial neural networks. Chemometrics and Intelligent Laboratory Systems, 98, 115-122.

[2] Bansode, N., & Sinha, P. (2016). Face shape classification based on region similarity, correlation and fractal dimensions. International Journal of Computer Science Issues, 13, 24-31.

[3] Cootes, T., Taylor, C., Cooper, D., & Graham, J. (1995). Active shape models: their training and application. Computer Vision and Image Understanding, 61, 38-59.

[4] Cootes, T. F, Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23, 681-685.

[5] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009) (pp. 248-255).

[6] Derrick, J. (2017). Is your face round, square, long, heart-shaped, or oval-shaped? URL: https://www.liveabout.com/is-your-face-round-square-long-heart-or-oval-shaped-345761.

[7] Fowler, M. (2015). The best haircut for every face shape. URL: https://theladylovescouture.com/the-best-haircut-for-every-face-shape.

[8] Fukushima, K., & Miyake, S. (1982). Neocognitron: A self-organising neural network model for a mechanism of visual pattern recognition. In Competition and Cooperation in Neural Nets (pp. 267-285).

[9] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 580-587).

[10] Gönen, M., & Alpaydın, E. (2011). Multiple kernel learning algorithms. Journal of Machine Learning Research, 12, 2211-2268.

[11] Grgic, M., Delac, K., & Grgic, S. (2011). SCface-surveillance cameras face database. Multimedia tools and applications, 51, 863-879.

[12] Hong, C. (2013). Powerful Hair Style (In Thai). Thailand: Nation International Edutainment PCL.

[13] Huang. G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). Labelled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst.

[14] Jain, V., & Learned-Miller, E. (2010). FDDB: A benchmark for face detection in unconstrained settings. Technical Report UM-CS-2010-009, University of Massachusetts, Amherst.

[15] Kang. W.-C., Fang. C., Wang, Z., & McAuley, J. (2017). Visually-aware fashion recommendation and design with generative image models. arXiv preprint arXiv:1711.02231.

[16] Li, L., So, J., Shin, H.-C., & Han, Y. (2013). An AAM-based face shape classification method is used for facial expression recognition. International Journal of Research in Engineering and Technology, 2, 164-168.

[17] Li, Y., Wu, F.-X., & Ngom, A. (2016). A review of machine learning principles for multi-view biological data integration. Briefings in Bioinformatics, (bbw113).

[18] Liang, Z., & Wang, F. (2016). Research on the personalised recommendation algorithm for hairdressers. Open Journal of Modelling and Simulation, 4, 102-108.

[19] Lienhart, R., & Maydt, J. (2002). An extended set of Haar-like features for rapid object detection. In Proceedings of the 2002 International Conference on Image Processing (ICIP'2002) (pp. 900-903).

[20] Lodkaew, T., Supsohmboon, W., Pasupa, K., & Loo, C. K. (2018). Fashion Finder: A system for locating online stores on Instagram from product images. In Proceedings of the 10th International Conference on Information Technology and Electrical Engineering (ICITEE 2018) (pp. 288-293).

[21] Lu, T. T. (2009). Fundamental limitations of semi-supervised learning. Master's thesis, University of Waterloo.

[22] Majtner, T., Yildirim-Yayilgan, S., & Hardeberg. J. Y. (2016). Combining deep learning and hand-crafted features for skin lesion classification. In Proceedings of the 6th International Conference on Image Processing Theory Tools and Applications (IPTA 2016) (pp. 1-6). IEEE.

[23] Mustafa, U. (2007). Face & eye detection. URL: https://www.mathworks.com/matlabcentral/fileexchange/13716-face---eye-detection.

[24] Ng, H.-W., Nguyen, V. D., Vonikakis, V., & Winkler, S. (2015). Deep learning for emotion recognition on small datasets using transfer learning. In Proceedings of the ACM International Conference on Multimodal Interaction (ICMI'2015) (pp. 443-449).

[25] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. In Proceedings of the 2015 British Machine Vision Conference (BMVC 2015) (pp. 1-12).

[26] Pasupa, K., Hussain, Z., Shawe-Taylor, J., & Willett, P. (2013). Drug screening with elastic-net multiple kernel learning. In Proceedings of the 13th International Conference on Bioinformatics and Bioengineering (BIBE'2013) (pp. 1-5).

[27] Pasupa, K., & Sunhem, W. (2016). A comparison between shallow and deep architecture classifiers on a small dataset. In Proceedings of the 8th International Conference on Information Technology and Electrical Engineering (ICITEE 2016) (pp. 1-6).

[28] Robnik-Šikonja, M., & Kononenko, I. (1997). An adaptation of Relief for attribute estimation in regression. In Proceedings of the 14th International Conference on Machine Learning (ICML1997) (pp. 296-304).

[29] Saragih, J. M., Lucey, S., & Cohn, J. F. (2009). Face alignment through subspace-constrained mean shifts. In Proceedings of the 12th International Conference on Computer Vision (ICCV'2009) (pp. 1034-1041).

[30] Sarakon, P., Charoenpong. T., & Charoensiriwath, S. (2014). Face shape classification from 3D human data using SVM. In Proceedings of the 7th Biomedical Engineering International Conference (BMEICON 2014) (pp. 1-5).

[31] Sharif Razavian, A., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN features off-the-shelf: an astounding baseline for recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 806-813).

[32] Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, L., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging, 35, 1285-1298.

[33] Siegel, S., & Castellan, N. J. (1988). Nonparametric statistics for the behavioural sciences (2nd ed.). McGraw-Hill.

[34] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[35] Sunhem, W., & Pasupa, K. (2016). An approach to face shape classification for hairstyle recommendation. In Proceedings of the 8th International Conference on Advanced Computational Intelligence (ICACI 2016) (pp. 390-394).

[36] Tan, B., Zhang, Y., Pan, S. J., & Yang, Q. (2017). Distant domain transfer learning. In Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI 2017) (pp. 2604-2610).

[37] Wang, N., Li, S., Gupta, A., & Yeung, D.-Y. (2015). Transferring rich feature hierarchies for robust visual tracking. arXiv preprint arXiv:1501.04587.

[38] Wang, P., Li, Z., Hou, Y., & Li, W. (2016). Combining convnets with hand-crafted features for action recognition based on an HMM-SVM classifier. arXiv preprint arXiv:1602.00749.

[39] Wu, S., Chen, Y.-C., Li, X., Wu, A.-C., You, J.-J., & Zheng, W.-S. (2016). An enhanced deep feature representation for person re-identification. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV 2016) (pp. 1-8).

[40] ZergNet (2017). Find the perfect cut for your face shape. URL: http://www.instyle.com/hair/find-perfect-cut-your-face-shape.

[41] Zhang, L., Zhang, D., Sun, M.-M., & Chen, F.-M. (2017). Facial beauty analysis based on geometric features: Toward attractiveness assessment application. Expert Systems with Applications, 82, 252-265.

[42] Zheng-Peng, W., & Zhang, X.-G. (2011). Elastic multiple kernel learning. Acta Automatica Sinica, 37, 693-699.

[43] Zhong, Y., Sullivan, J., & Li, H. (2016). Face attribute prediction using off-the-shelf CNN features. In Proceedings of the International Conference on Biometrics (ICB'2016) (pp. 1-7).