2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Facial Shape-Based Eyeglass Recommendation Using Convolutional Neural Networks

Rakib Hossain Rifat
Dept. of Computer Science and Engineering
BRAC University
Mohakhali, Dhaka-1212, Bangladesh
rakib.hossain.rifat@g.bracu.ac.bd

Sunzida Siddique *
Dept. of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
sunzida15-9667@diu.edu.bd

Laxmi Rani Das *
Dept. of Computer Science and Telecommunication Engineering
Noakhali Science and Technology University
Sonapur, Noakhali-3814, Bangladesh
laxmi0113@student.nstu.edu.bd

Mohd Ariful Haque
Dept. of Cyber-Physical Systems
Clark Atlanta University
Atlanta, Georgia
mohdariful.haque@students.cau.edu

***Abstract*—Eyeglasses are not only used to protect our vision and prevent dust from getting into our eyes. Additionally, glass that fits properly can give a person an elegant appearance. However, people often find it difficult to choose eyeglasses that fit their face shape; to address this issue, we have proposed a novel architecture in this paper. In order to do this, we created a pipeline that can recommend eyeglasses based on the form of the eyes using multiple transfer learning architecture to predict the face shape from a given image. We utilized InceptionV4 [17], InceptionV3 [18], Vit Small [12], DenseNet121 [10], ResNet50 [9], and VGG16 [16] to predict the facial shape from the image and achieve a test accuracy of 75%. We used 5500 photos with five different face shapes (Heart, Oblong, Oval, Round, Square) for this experiment, and two distinct datasets were gathered from Kaggle [2] and GitHub [1]. By simply uploading the photograph to our recommendation system, our proposed solution can assist users in selecting the appropriate eyewear.**

***Index Terms*—InceptionV4, InceptionV3, Vit Small, DenseNet121, ResNet50, VGG16, Trasnfer Learning, CNN, Eye Glass**

I. INTRODUCTION

A facial shape-based eyeglass suggestion system might assist the eyewear industry. Finding glasses that accentuate a person's facial traits has long been a challenge for consumers and eyeglass manufacturers. Eyewear has the dual function of enhancing one's vision and allowing one to exhibit a unique sense of style in a variety of settings. Face analysis is used by this suggestion system to assist customers in making the most appropriate frame selection, which in turn boosts both the customers' sense of style and self-confidence. Many different pairs are tried on by the customer, making it difficult for the sales associate to provide objective feedback.

A new eyeglass recommendation algorithm using Convolutional Neural Networks (CNNs) and face shape is presented in this paper. Convolutional Neural Networks (CNNs), deep learning algorithms, recommend eyeglasses based on facial features. This paper introduces a novel strategy for improving Convolutional Neural Networks (CNNs) in a domain. Careful processing and analysis of face photos allow for the proper selection of eyeglasses that match an individual's facial characteristics. This revolutionary technology optimizes eyewear selection by delivering personalized recommendations. Facial recognition, tracking, and verification require precise identification of individuals based on their facial features, especially form [7].

According to [4], the relationship between an object's shape and geometric structure is intricate. Understanding identification and user preferences is vital as a facial form is impacted by external and internal factors [14]. In the next parts of this study, we will assess facial forms to offer tailored eyewear suggestions, transforming the choosing process with powerful algorithms. This project aims to improve the user's experience and solve the problem of finding eyeglasses that improve visual acuity and match a person's face.

In the following parts, we will evaluate facial forms in order to make tailored recommendations for eyeglasses, thereby changing the selection process through the use of complex algorithms in order to improve the user experience.

II. LITERATURE REVIEW

This literature review explores various studies on analysing facial shapes and eyeglass try-ons. Some reviews are below-

The study by Young et. al [20] explored Convolutional Neural Networks for virtual glasses try-on. They employed the Viola-Jones algorithm on Android, using a dataset of 1300 human face photos, with a focus on oval and square faces. An 11-user test revealed only 3 successful frame attachments, resulting in a 27% satisfaction rate. The study highlights the need for improving accuracy and real-time face shape identification in future investigations.

Zafar et al. [21] used a data-driven approach for facial shape and eyeglass style classification with 99% accuracy for shape

*These authors contributed equally to this work.

extraction and 90% accuracy for classification using a nearest-neighbour method. A user study also showed an 82% agreement between recommended eyeglass frames and participants' preferences, confirming the efficacy of their methodology in personalized eyeglass recommendations.

Rohan et al [19] propose using Inception v3 to classify human faces into distinct shape categories. The study compares its performance with traditional classifiers (LDA, SVM, MLP, and KNN) on a dataset of 500 female celebrity images. Results show superior accuracy (98.0% to 100%) for training and overall accuracy (84.4% to 84.8%) compared to traditional classifiers, highlighting CNNs' effectiveness in face shape classification without manual feature selection. Dataset expansion is recommended for further model enhancement.

Deniz et al [6] introduced a live-video eyeglasses selection system based on computer vision techniques. It includes face and eye detection modules, improving eye detection accuracy compared to the Viola-Jones detector. Future enhancements are suggested, such as zoom and mirror-like spectacles adaptation.

Xiaoling et al [8] proposed a recommendation system for eyeglasses based on individual facial characteristics. Using a probabilistic graphical model, iGlasses captures dependencies between facial and frame attributes. Outperforming baseline methods on a dataset of face photos and eyeglass product images, iGlasses provides personalized recommendations, enhancing the shopping experience and increasing user satisfaction.

Another study Roha et.al [15] presents a computer-aided system that utilizes image processing and deep learning techniques to classify facial shapes into heart, oval, square, oblong, and circle types, achieving an 82% accuracy. While pre-defined features and classification methods like KNN, LDA, SVMLIN, and MLP are mentioned, the paper suggests incorporating convolutional neural networks (CNNs) to improve accuracy and generalize facial shape recognition across different contexts.

Mehta et al [14] propose a face shape classification method with 70% accuracy. The method uses facial landmarks and a classifier algorithm to categorize front-face images into heart, oblong, oval, round, and square shapes, outperforming existing approaches in landmark requirements and computational complexity. The study discusses the use of Active Shape Model (ASM), clustering algorithms, and support vector machines (SVM), highlighting the need for improvements in predicting other face shapes and acknowledging dataset limitations.

Duan et. al [7], proposes an algorithm utilizing M-RetinaFace for face alignment, EfficientNet with an attention mechanism for feature extraction, and a bilinear pooling layer for classification. With a dataset of 5500 images, the algorithm achieves an accuracy of 89.8% in classifying five face-shape categories. The combination of alignment, attention, and bilinear network enhances feature extraction, emphasizing the significance of accurate face shape classification in applications like face recognition and personalized recommendations.

The paper by Huang et al [11] presents a computer vision-based system using a webcam for virtual eyeglass try-on. The system tries on glasses via hand gestures with real-time images, limitations in real-time processing, and head rotation. It utilizes facial detection, eye-corner localization, fingertip detection, Harr-like features, and a cascaded AdaBoost classifier while acknowledging the limitations of non-frontal views and creating a realistic glasses image database using Maya.

Bansode et al [4] propose a novel approach for face shape classification using facial region similarity, correlation coefficient, and fractal dimensions. Evaluation on three datasets with 910 face images shows ellipse shape as the majority (78.73%), with other categories ranging from 1.76% to 10.23%. The approach demonstrates effective categorization based on geometric similarities.

Zhao et. al [22]proposed a facial attractiveness and achieves a high correlation (Pearson coefficient of 0.862) between machine predictions and human ratings, indicating a strong relationship between predicted and subjective attractiveness scores. The system utilizes 600 face images with positive and neutral expressions, demonstrating its potential in applications like facial beauty ranking, plastic surgery, and the entertainment industry.

The proposed Liu et. al [13] paper achieves 94% recognition accuracy. Comparisons with PCA and ELR methods confirm superior performance. Notably, in the Yale database, the recognition rate reaches 100% due to smaller, better samples. While the collected eyeglasses face dataset is comprehensive, it may not cover all styles and wearer appearances. They suggested exploring advanced feature extraction techniques, like deep learning architectures.

In the paper of Chu et al [5] the main contribution is the development of a 3D parametric human face modelling approach. They propose a framework that allows for capturing and modelling the unique characteristics of individuals' faces. The paper shows that the Kriging model with first-order regression achieves higher accuracy than linear regression. However, the limitations of the parametric modelling scheme are not discussed. Future work includes hair modelling and physical modelling techniques for more realistic face models.

III. DATASET

A. Data Collection

We have collected the training dataset from two different sources. The dataset1 [2] from Kaggle contains a total of 5000 images, and there are 5 classes in the dataset. And dataset2 [1] contains a total of 500 images of 5 classes. In order to increase the volume of the dataset we have mixed both datasets according to the class. Following Table I provides a short summary of the dataset.

Our final data doesn't contain any class imbalance, each class contains 20% of the whole dataset. A figure shown in Fig 1for Dataset Representation.

For training purposes, the dataset was divided into three splits: train, validation, and test (80:10:10). The distribution of data within each split is as follows Table II:

TABLE I
DATA DIST

Shape	Dataset 1	Dataset 2	Total
Heart	1000	100	1100
Oblong	1000	100	1100
Oval	1000	100	1100
Round	1000	100	1100
Square	1000	100	1100

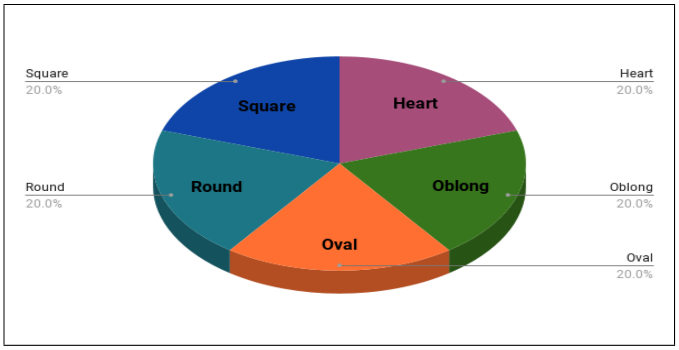


Fig. 1. Representation of Dataset

B. Data Preprocessing

The appearance and maintenance of one's physical appearance are contingent upon the individual's facial characteristics. Heart-shaped faces characterized by their feminine and aesthetically pleasing appearance exhibit a wider forehead and a pointed chin. Despite being longer and exhibiting a symmetrical proportion from the forehead to the chin, individuals with oblong faces possess a wider range of styling alternatives. Oval faces possess inherent qualities of balance and adaptability. Soft round faces have a greater degree of softness compared to their counterparts with powerful square faces.

The data preprocessing step involves transforming the raw images in the dataset to a format suitable for training and testing the Facial Shape Recognition system. This is achieved using a set of image transformations applied to both the training and testing datasets. The following transformations are applied:

1) **Conversion to tensor format:** The images are converted to tensors, which are numerical representations suitable for processing by the machine learning model.
2) **Resizing:** The images are resized to a specific height and width (224x224) to ensure uniformity in the input dimensions.
3) **Random horizontal and vertical flips:** The images are randomly flipped horizontally or vertically to introduce variations and augment the dataset.
4) **Random rotation:** The images are randomly rotated within a specified range (-10 to 10 degrees) to further augment the dataset and improve model robustness.
5) **Normalization:** The pixel values of the images are normalized using the mean and standard deviation values of imagenet (0.485, 0.456, 0.406) and (0.229, 0.224, 0.225), respectively. This step standardizes the pixel values, facilitating better convergence and training stability.

TABLE II
DATASET SPLITING

Train	Validation	Test
4400	550	550

For the testing dataset, similar transformations are applied, except for the random flips and rotations. The testing dataset undergoes resizing and normalization to ensure compatibility with the trained model during the evaluation phase.

IV. METHODOLOGY

Facial Shape Recognition in computer vision aims to accurately classify and analyze facial shapes from images. We merge datasets, create subsets with rigorous preprocessing, and train state-of-the-art models using optimized algorithms and loss functions. Evaluation metrics such as accuracy, precision, recall, and F1 score guide model selection. Comparative analyses emphasize the uniqueness and effectiveness of our approach.

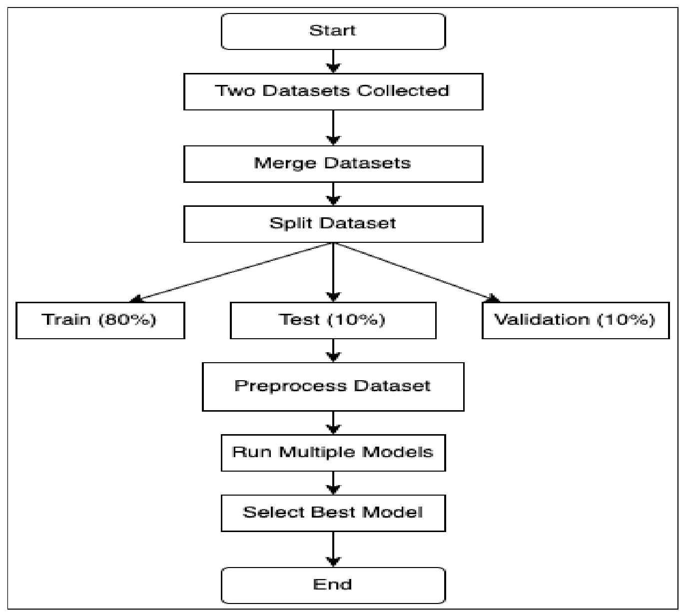


Fig. 2. Methodology for Training

1) **Data Collection:** Two distinct datasets, Kaggle [1] and GitHub [2], containing relevant facial images, are sourced for facial shape recognition research.
2) **Dataset Merging:** The collected datasets are merged to create a unified dataset for model training and evaluation.
3) **Data Splitting:** The merged dataset is divided into train, test, and validation subsets for specific data allocation in model training, testing, and unbiased evaluation.
4) **Data Preprocessing:** A rigorous pipeline including image resizing, pixel value normalization, and augmentation ensures standardized, high-quality, and diverse data, improving model performance.

5) **Model Training:** Multiple state-of-the-art models (InceptionV3 [18], InceptionV4 [17], ViT Small [12], DenseNet121 [10], ResNet50 [9], and VGG16 [16]) are trained on the preprocessed dataset with optimized algorithms and loss functions to facilitate convergence and generalization.
6) **Model Evaluation:** Trained models are assessed using robust metrics like accuracy, precision, recall, and F1 score. InceptionV4 excels in facial shape identification, outperforming other models. InceptionV3 also performs well, while DenseNet121 and ResNet50 achieve moderate accuracy, and Vit Small and VGG16 show relatively lower performance. A separate validation set ensures reliable evaluation of generalization across diverse datasets and unseen examples.
7) **Model Selection:** InceptionV4 is chosen as the optimal choice for the Facial Shape Recognition system, ensuring effectiveness and reliability in real-world applications.

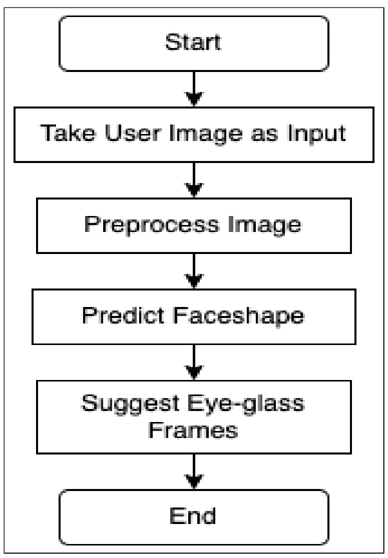


Fig. 3. Procedure in Website

The website allows easy photo upload and uses advanced preprocessing for precise facial shape analysis. The specialized machine learning model accurately classifies facial shapes (round, square, oval, heart, or oblong). Users receive glass frame recommendations based on their facial shape from a matching dataset of frames.

This methodology outlines a precise and rigorous process for developing and evaluating a Facial Shape Recognition system. It utilizes customized techniques, algorithms, and evaluation metrics to address specific research objectives, making a significant contribution to the field.

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. *Experimental Setup*

For our research, we use Google Colab GPU to employ deep learning methodologies. To be specific, we used the power of the Tesla T4 GPU for deep learning, while meticulously documenting all results in WandB. Our deep learning models were constructed using Python 3, PyTorch, and Poutyne. Additionally, we incorporated poutyne as a training wrapper, complemented by PyTorch Image Models (timm). To address data set balance, we leveraged the combination of deep learning models and imbalance learned in this experiment. In our models, we employed fixed hyperparameters, which are as follows in Table III:

TABLE III
HYPERPARAMETER VALUES

Parameter	Value
imgWidth, imgHeight	224, 224
Epochs	150
Batch size	32
Learning Rate	0.001
Mean	(0.485, 0.456, 0.406)
STD	(0.229, 0.224, 0.225)
Criterion	CrossEntropyLoss
Optimizer	Adam

B. *Results*

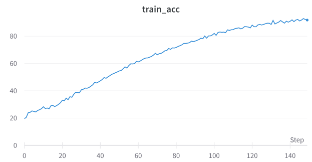
We evaluated the performance of several Convolutional Neural Network (CNN) models, namely InceptionV3, InceptionV4, ViT Small, DenseNet121, ResNet50, and VGG16, for facial shape-based eyeglass recommendation. In this section, we provide a detailed analysis of the performance comparison among these models.

Figure 4 showcases the performance metrics of our trained model, including test accuracy and loss values obtained during the evaluation phase. The InceptionV4 model achieved the best accuracy. Its test accuracy is 75.27% and a test loss is 0.85. During the training phase, the InceptionV4 model achieved a train accuracy of 91.77% and a train loss of 0.23. On the validation dataset, the InceptionV4 model attained a validation accuracy of 68.18% and a validation loss of 1.19. Although the validation accuracy is lower compared to the training accuracy, the model's performance showcases its capability to generalize and accurately classify instances on new and unseen data.

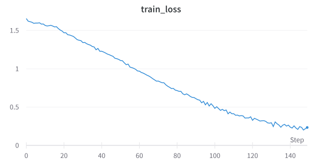
Overall, the evaluation results 4 demonstrate the effectiveness of the InceptionV4 model in facial shape-based eyeglass recommendation as it shows the best accuracy for our Model. Their evaluation figure is shown in [4(a), 4(b)], and [4(c), 4(d)]

A confusion matrix evaluates classification model performance through a table. After analyzing the confusion matrices, the Inception v4 model emerged as the top performer, as its diagonal elements indicated the highest number of correct predictions for each class.

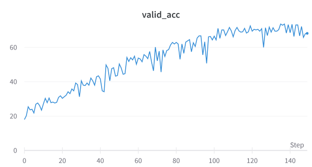
From the confusion matrices in Figure 5, we can see that the best-performing model Inception v4 has Correct predictions for heart class:79, Correct predictions for oblong class:96, Correct predictions for oval class:74, Correct predictions for round class:79, Correct predictions for square class:86 among the 114 data.



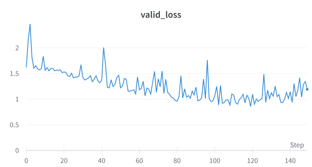
(a) Train Accuracy



(b) Train Loss

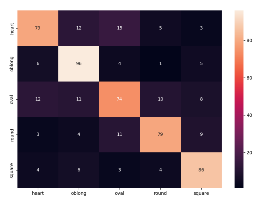


(c) validation Accuracy

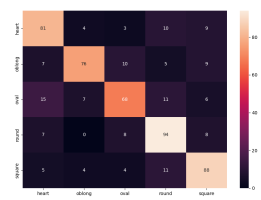


(d) Validation Loss

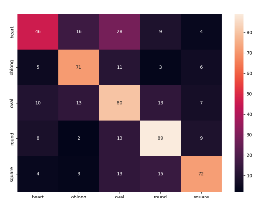
Fig. 4. Training, Testing, and Validation evaluation metrics of the InceptionV4 model



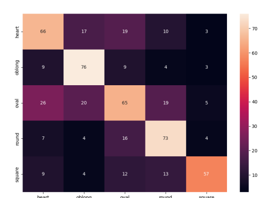
(a) Inception v4



(b) Inception v3



(c) DenseNet121



(d) ResNet50

Fig. 5. Confusion Matrix

By referring to Table IV, we can examine the results obtained from our models. The data highlights the superior performance of the Inception v4 models in comparison to the others. Inception v4 demonstrates the highest level of accuracy, reaching around 75.27%, and also performs well in terms of the F1-score of 80%. On the contrary, VGG16 exhibits the lowest performance, achieving an accuracy of only 22%.

VI. FRAMEWORK FOR RECOMMENDATION SYSTEM

In our project, we have successfully developed a model framework that optimizes the development process by saving valuable time. This framework empowers developers to focus on application tasks. By leveraging Python and FastAPI, an open-source Python framework, we have achieved this objective. The adherence of FastAPI to the model-view-controller architectural pattern significantly enhances the efficiency and

TABLE IV
EVALUATION SCORE OF CNN MODELS ON TEST DATASET

Class	CNN Models	Accuracy	Precision	Recall	F1-Score
Heart	Inception-v3	74%	70%	76%	73%
	Inception-v4	**75%**	76%	69%	72%
	ViT Small	39%	29%	30%	30%
	densenet121	65%	63%	45%	52%
Oblong	Inception-v3	74%	**84%**	71%	77%
	Inception-v4	75%	74%	**86%**	80%
	ViT Small	39%	36%	39%	38%
	densenet121	65%	68%	74%	71%
Oval	Inception-v3	74%	73%	64%	68%
	Inception-v4	75%	69%	64%	67%
	ViT Small	39%	46%	38%	42%
	densenet121	65%	55%	65%	60%
Round	Inception-v3	74%	72%	80%	76%
	Inception-v4	75%	80%	75%	77%
	ViT Small	39%	40%	51%	45%
	densenet121	65%	69%	74%	71%
Square	Inception-v3	74%	73%	69%	76%
	Inception-v4	75%	77%	83%	**80%**
	ViT Small	39%	47%	40%	43%
	densenet121	65%	73%	67%	65%

maintainability of our project [3]. As a result, we have established a robust and scalable solution for recommending eyeglasses based on face shape.

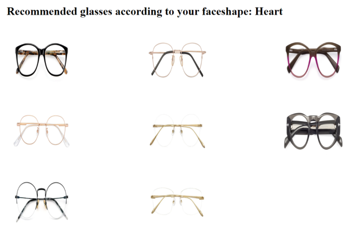
To illustrate the workflow, Figure 6 presents a user interface allowing users to upload their pictures. Subsequently, these pictures are automatically processed using our algorithmic approach and the designated model. The entire process operates the predefined algorithm. Within our research, this API serves as a testing API, employing our analytical algorithm methods in the background. Our primary objective is to enable users to upload pictures and automatically identify suitable eyeglasses based on their facial shape.

Fig. 6. Representation of Index page

Following the picture upload, a selection process determines the picture's suitability for further use, particularly in terms of eyeglass selection based on face shape criteria. Figure 7 demonstrates this process, whereby pictures meeting the specified criteria automatically identify suitable eyeglasses based on their facial shape.

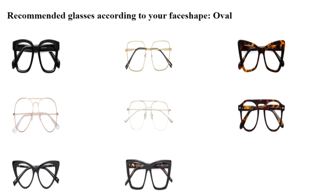
VII. CONCLUSION AND FUTURE WORK

This paper introduces a novel approach to eyeglass recommendations based on facial shape using Convolutional Neural Networks (CNNs). Our Facial Shape-Based Eyeglass Recommendation system leverages CNNs to identify and match ideal

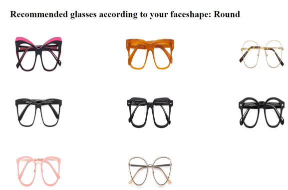


(a) Heart Recommendation

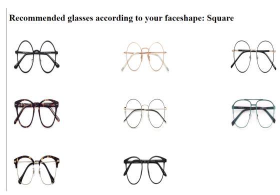
(b) Oblong Recommendation



(c) Oval Recommendation



(d) Round Recommendation



(e) Square Recommendation

Fig. 7. Framework-based Testing Result

eyeglass frames for personalized and aesthetically pleasing choices.

With InceptionV4 achieving an impressive test accuracy of approximately 75.27%, our system shows great potential to transform the eyewear industry by providing tailored recommendations. Moreover, our real-time application allows users to upload photos on a website for instant personalized eyeglass frame suggestions, enhancing the shopping experience and user satisfaction without in-store visits. By merging technology and fashion, we envision a future where individuals find their perfect pair of spectacles that celebrate their unique identity. With continued research dedication, our system can shape the future of eyewear selection, one personalized recommendation at a time.

The paper's limitation is the utilization of low-quality and online datasets, which consequently lead to relatively low model accuracy. In future, we aim to address these limitations by working with larger and more diverse datasets, exploring novel architectures, such as reinforcement learning, and focusing on improving the accuracy of the models.

References

[1] Github repository for face shape dataset. https://github.com/dsmlr/faceshape.

[2] Kaggle face shape dataset. https://www.kaggle.com/datasets/niten19/face-shape-dataset.

[3] Priya Bansal and Abdelkader Ouda. Study on integration of fastapi and machine learning for continuous authentication of behavioral biometrics. In *2022 International Symposium on Networks, Computers and Communications (ISNCC)*, pages 1–6. IEEE, 2022.

[4] NK Bansode and PK Sinha. Face shape classification based on region similarity, correlation and fractal dimensions. *International Journal of Computer Science Issues (IJCSI)*, 13(1):24, 2016.

[5] Chih-Hsing Chu, I-Jan Wang, Jeng-Bang Wang, and Yuan-Ping Luh. 3d parametric human face modeling for personalized product design: Eyeglasses frame design case. *Advanced Engineering Informatics*, 32:202–223, 2017.

[6] O Deniz, M Castrillon, J Lorenzo, M Hernandez, L Anton, and G Buenco. Eyewear selector. In *Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications-M2SFA2 2008*, 2008.

[7] Jiawei Duan, Xueping Su, Jie Ren, and Longsheng Xie. Face shape classification based on bilinear network with attention mechanism. In *Journal of Physics: Conference Series*, volume 2278, page 012041. IOP Publishing, 2022.

[8] Xiaoling Gu, Lidan Shou, Pai Peng, Ke Chen, Sai Wu, and Gang Chen. iglasses: A novel recommendation system for best-fit glasses. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 1109–1112, 2016.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[10] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[11] Wan-Yu Huang, Chaur-Heh Hsieh, and Jeng-Sheng Yeh. Vision-based virtual eyeglasses fitting system. In *2013 IEEE International Symposium on Consumer Electronics (ISCE)*, pages 45–46. IEEE, 2013.

[12] Seung Hoon Lee, Seunghyun Lee, and Byung Cheol Song. Vision transformer for small-size datasets. *arXiv preprint arXiv:2112.13492*, 2021.

[13] Li Liu, Yanfeng Sun, Baocai Yin, and Caifang Song. A novel nonuniform division strategy for wearing eyeglasses face recognition. In *2009 Fifth International Conference on Image and Graphics*, pages 907–911. IEEE, 2009.

[14] Ashwinee Mehta and T Mahmoud. Human face shape classification with machine learning. 2022.

[15] S Rohan, Suhas R Vittal, Neeraj H Gowda, and Dr. Priya R Sankapal. Face shape classifier using deep learning. 09, 2022.

[16] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[17] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

[18] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[19] Adonis Emmanuel Tio. Face shape classification using inception v3. *arXiv preprint arXiv:1911.07916*, 2019.

[20] Seo Eun Young, Friska Natalia, Sud Sudirman, and CS Ko. Eyeglasses frame selection based on oval face shape using convolutional neural network. 2019.

[21] Amir Zafar and Tiberiu Popa. Face and eye-ware classification using geometric features for a data-driven eye-ware recommendation system. 2016.

[22] Jian Zhao, Miao Zhang, Chen He, Xie Xie, and Jiaming Li. A novel facial attractiveness evaluation system based on face shape, facial structure features and skin. *Cognitive Neurodynamics*, 14:643–656, 2020.