Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Facial Beauty Standards Predictions Based on Machine Learning: A Comparative Analysis

Bareen Haval Sadiq[1*], Adnan M. Abdulazeez [2]
bareen.haval@dpu.edu.krd, adnan.mohsin@dpu.edu.krd
[1,2] IT Department, Technical College, Duhok Polytechnic University, Kurdistan Region/Iraq

Article Information

Submitted : 22 Jan 2024
Reviewed: 25 Jan 2024
Accepted : 10 Feb 2024

Keywords

Facial beauty standards, Random Forest, KNN, SVM, Decision Tree, Logistic Regression

Abstract

This study uses a variety of machine learning and classification methods to anticipate the Facial Beauty Standards. The Accuracy of five different models—Random Forest, Logistic Regression, Support Vector Machine (SVM), KNN, and decision tree—were used to analyses each one. There were noticeable differences in the models' performances. In particular, the Logistic Regression and SVM methods demonstrate almost perfect accuracy, followed closely by random forest and KNN. This study gives insight into how well different models perform in comparison and emphasizes the benefits and drawbacks of each in terms of predicting face beauty standards.

I. INTRODUCTION

Recent years have seen the convergence of face analysis and artificial intelligence to provide new avenues for study, particularly in the field of facial beauty analysis and prediction. The ability to quantify and assess face beauty objectively has several applications, such as developments in facial recognition technology and the cosmetics industry. Ancient Greek and Chinese scholars created notions such as the "golden ratio" and "three courts and five eyes" to measure face beauty. Recently, scientists have begun using modern statistics and machine learning methods to evaluate human looks. Many statistical and geometrical rules have been proposed, such as the "symmetrical face", the "average face", and the "golden ratio". In particular, machine learning models may be used to recover geometric data, such as the distance and ratio characteristics, and facial landmarks, which are then utilized to evaluate the facial beauty. These studies satisfy both the existence of facial geometric laws and human subjectivity by focusing on the facial organs—such as the mouth, eyes, nose, chin, eyebrows, and so on—and their relationships. For facial beauty analysis, texture features and color features [1] have been suggested in addition to geometric features. In the fields of computer vision and machine learning, there is increasing interest in the automatic human-like Facial Attractiveness Prediction (FAP) or Face Beauty Prediction (FBP). There will likely be a greater application case: facial image beautifying, social network recommendation systems, makeup evaluation, facial image enhancement, content-based face retrieval, cosmetic suggestion, planning aesthetic surgery, picture editing, and more [2], and [3]. Research in psychology revealed that people of different backgrounds all saw beauty in the same ways. Every day, billions of photos are taken and posted to social media and other online platforms in the age of mobile computing. It presents a fresh challenge to technology used for picture processing and analysis. Big data and high-performance computing hardware have recently led to the proposal of computational and data-driven methods for addressing these problems, including face recognition [4], facial emotion recognition [5], facial beauty analysis [6], and others. From the numerical descriptions offered in early explorations to the measurement-based standards used in aesthetic surgery, and from the theories put forth by contemporary psychologists to computer models for assessing facial attractiveness, the history of facial attractiveness research can be broken down into multiple phases [7]; [8] Complicated mental processes that vary depending on the person are part of the cognitive processes associated with facial attractiveness. Yet, it was discovered that people's assessments of face attractiveness were consistent for a certain historical period and cultural heritage [9] The results of studies from several disciplines complement one another and advance our knowledge of facial beauty cognition. These days, machine learning (ML) is used to all computational tasks where performance is improved and methods are developed [10] Learning from unbalanced data sets has emerged as a major machine learning challenge in recent years and is widely used in a variety of applications, including computer security, engineering, remote sensing, biology, computer security, and industry transformation. When the target variable is categorical in classification and keeps declining, supervised learning approaches such as band algorithms, regression, and classification are used [10]. A small number of attributes and many samples (tuples) make up machine learning datasets. The typical machine learning datasets are not exactly the same as microarray technology Support vector machines (SVMs) are supervised learning techniques used in categorization. The goal of this research article is to investigate and evaluate how well different machine learning models predict facial beauty criteria. The models that have been chosen represent a range of methodologies, each possessing distinct attributes and flexibility to accurately anticipate standards of facial beauty, it tests the performance of Random Forest, K-Nearest Neighbors (KNN), Decision Trees, Logistic Regression, and Support Vector Machines (SVM).

B. Literature review

Both linear and nonlinear machine learning techniques have seen a variety of applications recently. As a result, utilizing deep structure—a more adaptable subset of machine learnin. in 2020 integrated manifold learning and feature extraction [11]. In [12]the findings of the experiment, the prediction score of face attractiveness was raised when facial landmarks and primary facial traits were combined. [13] provide a two-branch architecture (REX-INCEP) that addresses the intricate high-level characteristics connected to the FBP issue by combining the architectures of two previously trained networks. To extract the most pertinent and discriminant features from an input face picture or face descriptor, present a multi-layer local discriminant embedding approach that incorporates feature selection as a primary step. With the help of a cascaded feature extraction and selection architecture that can turn noisy and weak descriptors into strong ones, the suggested framework enables the translation of any linear approach into a deep variation was presented in [14] [14] .

To predict face beauty, [15] introduce a network called Multi-input Multi-task Beauty Network (2M BeautyNet) and employ transfer learning. Gender recognition is an ancillary job in the experiment, with attractiveness prediction serving as the primary objective. [16] suggested approach makes use of continuous ratings throughout the whole range and is based on texture. Second, we modify and kernelize a linear Flexible Manifold Embedding strategy that already exists and is applicable to discrete classes in order to handle the propagation of actual scores. Both inductive and transudative contexts can employ the final model.

In [17] the goal is to develop a better representation of face information by distributing the significance across aspects through the investigation of a combined spatial-wise and channel-wise attention (SCA) block. The suggested network can predict face attractiveness more accurately than the state-of-the-art, according to experimental data. For Asian female face attractiveness prediction tasks, a CNN approach based on transfer learning is applied, which incorporates various channel characteristics, a more realistically distributed Large Scale Asian Female Beauty Dataset (LSAFBD) has been created in [18]. adopted a co-attention learning process to identify the importance of several face components and areas at the same time was presented in [19] It uses the SCUT-FBP5500 and Celeb A datasets for its tests. The outcomes demonstrate that the accuracy of the face attractiveness prediction is much increased by our co-attention learning approach. A conceptual classification of beautification is offered, along with a discussion of pertinent facial recognition scenarios and a review of relevant literature is presented in [20], along with unresolved problems and difficulties in the field, technical concerns and trade-offs of the assessed methodologies are also discussed. [21] provide a thorough and innovative approach for evaluating the attractiveness of the face that is based on the structural aspects of the skin, the facial structure, and the form of the face. The initial step in using face shape structural elements to facial attractiveness assessment is to classify face shapes. The facial structure and skin texture characteristics that indicate facial attractiveness are retrieved from the face picture dataset and fused after it has been separated into segments based on face shape.

To predict facial attractiveness, the machine learning method that performs the best in prediction is chosen from the face shape structural subgroups. [22]suggest utilizing face recognition data to do computer-aided facial diagnosis on a range of illnesses via deep transfer learning. Using a rather modest dataset, the computer-aided face diagnosis was carried out in the tests on a single illness (beta-thalassemia) as well as many disorders (hyperthyroidism, leprosy, Down syndrome, and beta-thalassemia). Rather of relying just on a single face descriptor, the suggested approach incorporates both geometric and deep feature-based graphs to provide a high-level representation of face pictures. It also enhances the discriminative power of graph-based score propagation techniques is discussed in [23]. In [24] the approach involves creating a face identification framework that uses 2D facial photos collected from several sources to create a 3D face mesh with 468 Media Pipe markers that can identify numerous faces in real time. [25] applies random forests (RF) to sets of multitemporal metrics that take seasonal within-class dynamics into account in order to conduct LULC classification also Robust change vector analysis (RCVA) is used for spectral change detection in order to identify changes that may not always result in a different class. [26]suggested combining the L1, L2, and Log-cosh regression-loss functions into an ensemble and then averaging them to produce a new composite cost function. By leveraging the distinctive qualities of every loss function, this method creates a cohesive framework that balances accuracy, flexibility, and outlier tolerance. [27]offers the facial image attractiveness classification network (FIAC-Net), a light deep convolutional neural network (LDCNN) with fewer parameters than pre-trained networks, enabling the deployment of deep neural networks on small hardware for the assessment of both genders' facial image attractiveness.

C. The proposed method

Figure 1 shows the broad framework of the suggested approach. The input data were first divided into sets for training and testing. After that, the training data were used to teach the classification algorithms how to estimate beauty standards in a face image.

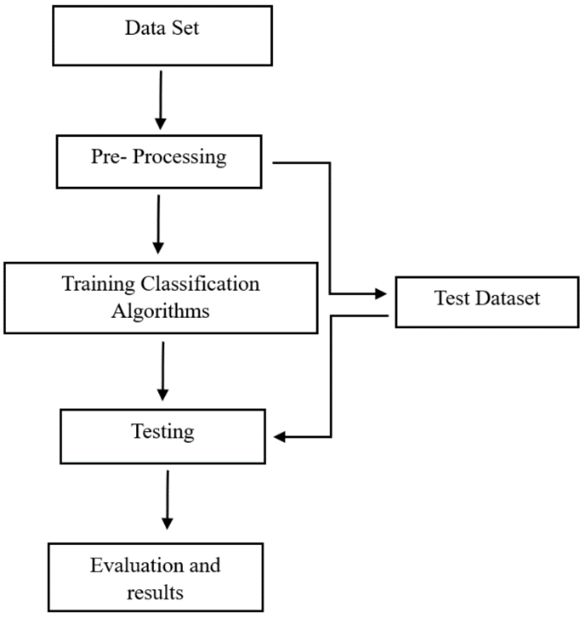

Figure 1. Data Flow of Proposed Models

D. Pre-processing

A powerful tool for machine learning, Keras' Image Data Generator allows for the quick loading, augmentation, and preparation of image data is used for the pre-processing [28]. It is Particularly useful for computer vision tasks. In order to save memory and enable real-time data augmentation, this module provides an adaptable method of preparing and enhancing picture datasets on the fly throughout the model training process. You may apply a variety of picture data transformations, including rotation, shifting, flipping, and zooming, normalization, and more, by using Image Data Generator. This augmentation can reduce overfitting and improve the model's ability to generalize to new data by broadening the training dataset. [29].

a. Random Forest

A collection of tree-structured classifiers with independently distributed random vectors that are all uniformly distributed make up the Random Forest classifier. At input x, each tree casts a unit vote for the class that is most popular [30]. An upper bound is extracted for Random Forests to obtain the generalization error in terms of two parameters: the interdependence of individual classifiers and the accuracy of the generated random vector, which is independent of the previous random vectors of the same distribution. A tree is then generated using the training test. The Random Forest flow chart is displayed in Figure 2. [29]. The Random Forest algorithm determines prediction $\hat{Y}$ for a given input X by averaging the forecasts of individual trees T:

$$\hat{Y} = \frac{1}{N} \sum_{i=1}^{N} T_i(X) \qquad (1)$$

Where N is the forest's tree count. The Random Forest model's hyperparameters, which include the number of trees, maximum depth, and minimum samples per leaf, are tuned to either maximize the coefficient of determination or minimize the mean squared error (MSE).

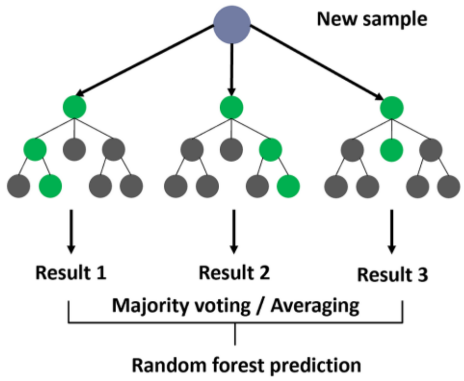

Figure 2. Random Forest general example [29]

b. DECISION TREE

A well-liked machine learning approach for classification and regression applications is the decision tree [31]. This approach belongs to the family of supervised learning algorithms called "tree-based models." This is a quick synopsis of machine learning using decision trees. A decision tree is a type of hierarchical tree structure in which a decision or testing on a characteristic is represented by each node, the test result is represented by each branch, and the ultimate decision or goal variable is represented by each leaf node [32].

c. K-NEAREST NEIGHBORS KNN

KNN is widely used in data identification and consistency for regression and classification. One supervised machine learning algorithm in the KNN family is KNN. It is a non-parametric approach to statistical analysis. A corresponding goal and output model are created once the input, which is the learning input in both scenarios, is taken from a training block. K-NN is a sort of memory-based learning as inferences are made directly from training instances. The neighbors are drawn from an object class that is well-known. The class is allocated to its single nearest neighbor if K = 1. When there is the smallest distance between two peers—known as the Euclidean distance—a straight line will always be created using a conventional clustering method [33]. The mean of its KNNs is the outcome of the KNN regression. The KNN method has some drawbacks, such as being dependent on local data alignment, not being the quickest, only working with a small number of inputs, and requiring homogenous features. The following are the equations for KNN: In Equation 2, the Euclidean equation (Ee) is illustrated [34].

$$\mathfrak{E}_{\mathfrak{e}} = \sqrt{\sum_{j=1}^{H}(\alpha_j - \beta_j)^2} \qquad (2)$$

Manhattan equation (Me) is presented in Eq.3

$$\mathfrak{M}_{\mathfrak{e}} = \sum_{j=1}^{H} |\alpha_j - \beta_j| \qquad (3)$$

Minkows equation (Mke) is presented in Eq. 4

$$\mathfrak{M}\mathfrak{K}_{\mathfrak{e}} = \left(\sum_{j=1}^{H}(|\alpha - \beta_j|)^p\right)^{1/p} \qquad (4)$$

d. Logistic Regression

A popular statistical technique in machine learning for binary classification issues is logistic regression. Logistic regression is not used for regression; rather, it is used for classification. It forecasts the likelihood that a given incident falls into a specific category. There are two potential outcomes in binary classification: 0 or 1, which are frequently associated with positive and negative classifications or, more accurately, with "not happening" and "happening." [35]. The logistic function, sometimes referred to as the sigmoid function, is used to a linear combination of input characteristics in the logistic regression model. The logistic function may be used to express probabilities since it translates every real-valued quantity to the interval between 0 and 1.

e. SVM

A method called support vector machines (SVMs) is particularly helpful for regression, classification, and general pattern recognition, respectively. Its ability to generalize when supervised training data is not required makes it a popular classifier. It is deemed appropriate since it is a generalizer, even with a high number of input vectors. A linear function f runs across the class midpoints for a dataset that is linearly separable, representing a line with the formula y = f(x) [36]

Figure 3. SVM General Example [37]

Is a statistical method that, when applied to one or more independent variables and a binary dependent variable, describes the connection as follows:

$$l = \log p\left(\frac{p_y}{1-p_y}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + + \beta_n x_n \quad (5)$$

E. Results and Discussion

The "Beauty Classification: Beautiful or Average" dataset, a comprehensive collection that can be used to train and evaluate algorithms for beauty classification, was utilized to test the models' performance. You can get it at Kaggle [38]. This dataset covers a wide range of facial features, emotions, and aesthetics and comprises of 300 testing pictures and 4000 properly annotated training shots. The training subset consists of a large variety of images with various **attributes** related to attractiveness. This dataset designates some aspects as being indicative of beauty, such as symmetry, flawless skin, attractive eyes, and distinct facial structures. Additionally, the dataset ensures inclusivity by incorporating a wide range of ages, ethnicities, and cultural backgrounds to provide models a comprehensive learning experience. In addition, the 300-image testing set provides a dependable platform for evaluating the accuracy and generality of the model in recognizing beauty classes within previously unseen data, which is of interest to academics and developers. This Kaggle-generated dataset is a helpful resource for improving computer vision algorithms that identify and categorize visual data's aesthetic aspects. Five distinct models were used: SVM, Random Forest, KNN, and Logistic Regression. Every model went through extensive testing and training, with accuracy ratings serving as the focal point of assessment criteria. The Random Forest algorithm performs very well with complicated and multidimensional datasets because of its ensemble learning methodology. Random Forest is an effective technique for capturing the subtleties of beauty standards since it generates reliable forecasts by combining the outputs of several decision trees. Its versatility and capacity to adjust to shifting fashions make it a strong competitor in our investigation. KNN finds patterns in the data by comparing the similarities between occurrences. This technique may be useful for forecasting beauty standards since it may be used to find attractiveness clusters and detect commonalities among aesthetically pleasant characteristics.

Decision Tree algorithms provide a clear window into the process of making decisions. Decision Trees show the hierarchy of elements impacting beauty standards by recursively splitting the data depending on attributes. Decision Trees are useful for comprehending the fundamental ideas guiding the sense of beauty because of their interpretability. A well-known algorithm in the field of binary classification, logistic regression offers a linear method for figuring out how features and results relate to one another. Logistic regression, a straightforward statistical technique, may bridge the gap between more sophisticated machine learning techniques and classic statistical methods by providing insightful information about the linear features of beauty criteria. Encouragement Recognized for its prowess in managing high-dimensional data, support vector machines can capture complex patterns while maintaining aesthetic standards. SVMs negotiate the intricacy of beauty perceptions by constructing a hyperplane that optimally separates several classes and pinpointing limits that capture the subtleties of attractiveness. When the Accuracy value is near to 1, it indicates that the dependent variable's variation is adequately explained by the model. When the Accuracy value is around zero, it suggests that the model may not be a good fit and is unable to clarify much of the variation. The Accuracy formula is:

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} * 100$$

Table 1 Accuracy Comparison for Proposed Models

No	Models	Accuracy
1	Random Forest	0.983
2	Logistic Regression	0.993
3	SVM	0.993
4	Decision Tree	0.87
5	KNN	0.946

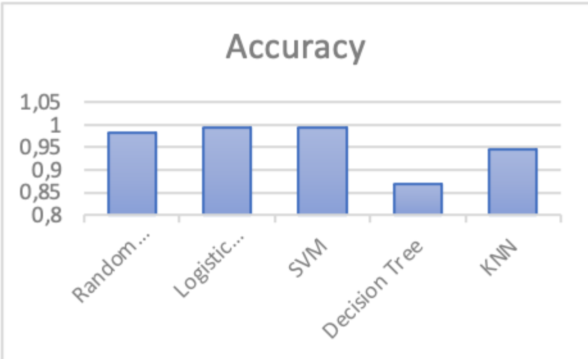

Figure 3. Accuracy Comparison for Proposed Models

The Random Forest model showed performance in predicting facial beauty standards prediction, approximately 98.3% of the variation in the dependent variable can be explained by the Random Forest model. Logistic Regression, a fundamental yet interpretable model, with an Accuracy of 0.993, the logistic regression model is able to account for 99.3% of the variation in the dependant variable. Additionally, the SVM model has a high amount of variance explained (99.3%) with an Accuracy of 0.993. SVM's excellent predicted accuracy can be attributed to its capacity to handle non-linear connections using kernel functions, which allowed it to identify complex patterns in the data. Table 1 and figure 4 illustrate the accuracy Comparison.

Approximately 87% of the variation in the dependent variable can be explained by the Decision Tree model. When viewed alongside the other models, this Accuracy value is lower. With an Accuracy of 0.946, the KNN model can account for around 94.6% of the variation in the dependent variable.

F. Conclusion

The combination of face analysis with artificial intelligence has created new and interesting research opportunities, especially in the area of facial beauty analysis and prediction. This combination of technologies has found uses in many different sectors, from advances in the cosmetics business to facial recognition technology. Modern methods use machine learning techniques and statistical standards to objectively evaluate beauty, whereas historical viewpoints on face attractiveness, such as the "golden ratio" and ancient Greek and Chinese philosophies, set the foundation. The accuracy and impartiality of face attractiveness assessments have been greatly improved by the inclusion of facial landmarks, geometric data, texture characteristics, and colour variables in machine learning models. Notably, several models have shown differing degrees of effectiveness in predicting face beauty standards, including the Random Forest, Logistic Regression, SVM, Decision Tree, and KNN. Particularly impressive explanatory power was demonstrated by the Logistic Regression, and SVM models, which together accounted for 99% of the differences in the variable being studied. All things considered, the nexus between AI and face recognition presents a potentially fruitful path toward objective evaluations of beauty, and the variety of machine learning models available gives practitioners and academics a toolset to customize their strategies according to particular needs. In order to guarantee the practical usability and relevance of these breakthroughs in real-world contexts, it is imperative that contextual factors be balanced with technical measures like Accuracy as this area continues to progress.

G. References

[1] L. Xu, J. Xiang, and X. Yuan, “CRNet: classification and regression neural network for facial beauty prediction,” in *Pacific Rim Conference on Multimedia*, Springer, 2018, pp. 661–671.

[2] P. Vatiwutipong, S. Vachmanus, T. Noraset, and S. Tuarob, “Artificial Intelligence in Cosmetic Dermatology: A Systematic Literature Review,” *IEEE Access*, 2023.

[3] J. Saeed and A. M. Abdulazeez, “Facial beauty prediction and analysis based on deep convolutional neural network: a review,” *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 1–12, 2021.

[4] K. Cao, K. Choi, H. Jung, and L. Duan, “Deep learning for facial beauty prediction,” *Information*, vol. 11, no. 8, p. 391, 2020.

[5] P. A. Martin Cervantes, N. Rueda Lopez, and S. Cruz Rambaud, “Life expectancy at birth in Europe: An econometric approach based on Random Forests methodology,” *Sustainability*, vol. 12, no. 1, p. 413, 2020.

[6] D. M. Abdulqader, A. M. Abdulazeez, and D. Q. Zeebaree, “Machine learning supervised algorithms of gene selection: A review,” *Mach Learn*, vol. 62, no. 03, pp. 233–244, 2020.

[7] M. A. Sulaiman, “Evaluating data mining classification methods performance in internet of things applications,” *Journal of Soft Computing and Data Mining*, vol. 1, no. 2, pp. 11–25, 2020.

[8] A. M. Abdulazeez, D. Q. Zeebaree, D. A. Zebari, and T. H. Hameed, “Leaf Identification Based on Shape, Color, Texture and Vines Using Probabilistic Neural Network,” *Computación y Sistemas*, vol. 25, no. 3, pp. 617–631, 2021.

[9] T. J Iyer, R. Nersisson, Z. Zhuang, A. N. Joseph Raj, and I. Refayee, “Machine learning-based facial beauty prediction and analysis of frontal facial images using facial landmarks and traditional image descriptors,” *Comput Intell Neurosci*, vol. 2021, 2021.

[10] F. Bougourzi, F. Dornaika, and A. Taleb-Ahmed, “Deep learning based face beauty prediction via dynamic robust losses and ensemble regression,” *Knowl Based Syst*, vol. 242, p. 108246, 2022.

[11] F. Dornaika, A. Moujahid, K. Wang, and X. Feng, “Efficient deep discriminant embedding: Application to face beauty prediction and classification,” *Eng Appl Artif Intell*, vol. 95, p. 103831, 2020.

[12] J. Zhao, M. Zhang, C. He, X. Xie, and J. Li, “A novel facial attractiveness evaluation system based on face shape, facial structure features and skin,” *Cogn Neurodyn*, vol. 14, pp. 643–656, 2020.

[13] F. Bougourzi, F. Dornaika, and A. Taleb-Ahmed, “Deep learning based face beauty prediction via dynamic robust losses and ensemble regression,” *Knowl Based Syst*, vol. 242, p. 108246, 2022.

[14] S. Shi, “Improving facial attractiveness prediction via co-attention learning in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019.” IEEE.

[15] J. Gan *et al.*, “2M BeautyNet: Facial beauty prediction based on multi-task transfer learning,” *IEEE Access*, vol. 8, pp. 20245–20256, 2020.

[16] F. Dornaika, K. Wang, I. Arganda-Carreras, A. Elorza, and A. Moujahid, “Toward graph-based semi-supervised face beauty prediction,” *Expert Syst Appl*, vol. 142, p. 112990, 2020.

[17] K. Cao, K. Choi, H. Jung, and L. Duan, “Deep learning for facial beauty prediction,” *Information*, vol. 11, no. 8, p. 391, 2020.

[18] J. Saeed and A. M. Abdulazeez, “Facial beauty prediction and analysis based on deep convolutional neural network: a review,” *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 1–12, 2021.

[19] S. Shi, F. Gao, X. Meng, X. Xu, and J. Zhu, “Improving facial attractiveness prediction via co-attention learning,” in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 4045–4049.

[20] C. Rathgeb, A. Dantcheva, and C. Busch, “Impact and detection of facial beautification in face recognition: An overview,” *IEEE Access*, vol. 7, pp. 152667–152678, 2019.

[21] J. Zhao, M. Zhang, C. He, X. Xie, and J. Li, “A novel facial attractiveness evaluation system based on face shape, facial structure features and skin,” *Cogn Neurodyn*, vol. 14, pp. 643–656, 2020.

[22] B. Jin, L. Cruz, and N. Gonçalves, “Deep facial diagnosis: deep transfer learning from face recognition to facial diagnosis,” *IEEE Access*, vol. 8, pp. 123649–123661, 2020.

[23] F. Dornaika and A. Moujahid, “Multi-view graph fusion for semi-supervised learning: application to image-based face beauty prediction,” *Algorithms*, vol. 15, no. 6, p. 207, 2022.

[24] G. Sanil, K. Prakash, S. Prabhu, V. Nayak, and S. Sengupta, “2D-3D Facial Image Analysis for Identification of Facial Features Using Machine Learning Algorithms with Hyper-parameter Optimization for Forensics Applications,” *IEEE Access*, 2023.

[25] F. Thonfeld, S. Steinbach, J. Muro, and F. Kirimi, “Long-term land use/land cover change assessment of the Kilombero catchment in Tanzania using

random forest classification and robust change vector analysis," *Remote Sens (Basel)*, vol. 12, no. 7, p. 1057, 2020.
[26] J. N. Saeed, A. M. Abdulazeez, and D. A. Ibrahim, "Automatic Facial Aesthetic Prediction Based on Deep Learning with Loss Ensembles," *Applied Sciences*, vol. 13, no. 17, p. 9728, 2023.
[27] J. N. Saeed, A. M. Abdulazeez, and D. A. Ibrahim, "FIAC-Net: Facial image attractiveness classification based on light deep convolutional neural network," in *2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA)*, IEEE, 2022, pp. 1–6.
[28] A. A. Salih and A. M. Abdulazeez, "Evaluation of classification algorithms for intrusion detection system: A review," *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 31–40, 2021.
[29] N. Mohammad, A. M. Muad, R. Ahmad, and M. Y. P. M. Yusof, "Accuracy of advanced deep learning with tensorflow and keras for classifying teeth developmental stages in digital panoramic imaging," *BMC Med Imaging*, vol. 22, no. 1, p. 66, 2022.
[30] D. Wang and A.-X. Zhu, "Soil mapping based on the integration of the similarity-based approach and random forests," *Land (Basel)*, vol. 9, no. 6, p. 174, 2020.
[31] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, 2021.
[32] D. M. Ahmed, A. M. Abdulazeez, D. Q. Zeebaree, and F. Y. H. Ahmed, "Predicting university's students performance based on machine learning techniques," in *2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS)*, IEEE, 2021, pp. 276–281.
[33] N. Deepa *et al.*, "A survey on blockchain for big data: Approaches, opportunities, and future directions," *Future Generation Computer Systems*, vol. 131, pp. 209–226, 2022.
[34] T. Alquthami, M. Zulfiqar, M. Kamran, A. H. Milyani, and M. B. Rasheed, "A performance comparison of machine learning algorithms for load forecasting in smart grid," *IEEE Access*, vol. 10, pp. 48419–48433, 2022.
[35] X. Song, X. Liu, F. Liu, and C. Wang, "Comparison of machine learning and logistic regression models in predicting acute kidney injury: A systematic review and meta-analysis," *Int J Med Inform*, vol. 151, p. 104484, 2021.
[36] S. H. Haji, A. M. Abdulazeez, D. Q. Zeebaree, F. Y. H. Ahmed, and D. A. Zebari, "The impact of different data mining classification techniques in different datasets," in *2021 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, IEEE, 2021, pp. 1–6.
[37] J. Cardoso-Fernandes, A. C. Teodoro, A. Lima, and E. Roda-Robles, "Semi-automatization of support vector machines to map lithium (Li) bearing pegmatites," *Remote Sens (Basel)*, vol. 12, no. 14, p. 2319, 2020.
[38] "https://www.kaggle.com/datasets".