Data and Metadata. 2025; 4:658
doi: 10.56294/dm2025658

REVIEW

Systematic Review: Recent Advancements in Deep Learning Techniques for Facial Feature Recognition

Revisión sistemática: avances recientes en técnicas de aprendizaje profundo para el reconocimiento de rasgos facials

Srinivas Adapa[1], Vamsidhar Enireddy[1]

[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation. Andhra Pradesh, 522302, India.

Cite as: Adapa S, Enireddy V. Systematic Review: Recent Advancements in Deep Learning Techniques for Facial Feature Recognition. Data and Metadata. 2025; 4:658. https://doi.org/10.56294/dm2025658

Submitted: 11-02-2024 **Revised:** 16-07-2024 **Accepted:** 11-02-2025 **Published:** 12-02-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding Author: Srinivas Adapa

ABSTRACT

Deep Learning is a rapidly evolving field with critical contributions to various domains including security, healthcare, and human — computer interaction, etc. It reviews the significant developments in the area of facial recognition using deep learning techniques. It explains deep learning models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), and Generative Adversarial Networks (GANs), as well as hybrid models and transfer learning uses. It also addresses technical, ethical, and legal challenges that arise for facial analysis systems and emphasizes the need for real-time processing, multi-modal systems, and robust algorithms to improve the technical accuracy and fairness of facial analysis.

Keywords: Convolutional Neural Networks; Long Short-Term Memory Networks; Generative Adversarial Networks; Hybrid Deep Learning Approaches; Human-Computer Interaction; Multi-modal Systems.

RESUMEN

El aprendizaje profundo es un campo en rápida evolución con contribuciones críticas a diversos dominios, como la seguridad, la atención sanitaria y la interacción persona-ordenador, etc. Se revisan los avances significativos en el área del reconocimiento facial mediante técnicas de aprendizaje profundo. Explica los modelos de aprendizaje profundo, como las redes neuronales convolucionales (CNN), las redes neuronales recurrentes (RNN), las redes de memoria a largo plazo (LSTM) y las redes generativas adversariales (GAN), así como los modelos híbridos y los usos del aprendizaje por transferencia. También aborda los problemas técnicos, éticos y jurídicos que plantean los sistemas de análisis facial y hace hincapié en la necesidad de un procesamiento en tiempo real, de sistemas multimodales y de algoritmos robustos para mejorar la precisión técnica y la imparcialidad del análisis facial.

Palabras clave: Redes Neuronales Convolucionales; Redes de Memoria a Largo Plazo; Redes Generativas Adversariales; Enfoques Híbridos de Aprendizaje Profundo; Interacción Persona-Ordenador; Sistemas Multimodales.

INTRODUCTION

Humans have a natural attraction to facial beauty. Awareness of beauty is becoming increasingly crucial in medical settings, as the demand for cosmetic surgery has increased significantly in recent years. The physical beauty of faces influences a wide range of social outcomes, including partner and employment decisions.[1] The cosmetics industry has developed a wide range of products to improve various body parts, including hair, skin, eyes, brows, and lips. It should come as no surprise that facial features have been studied for many years in psychology and other scientific fields.[2] In order to extract images from cameras for processing and analysis in a way that is similar to the natural human vision system, computer vision systems have been used extensively in recent years.[3]

Deep learning techniques have significantly improved facial recognition technology, resulting in accurate and robust systems for face recognition across various environments. Data-driven techniques that can learn complicated representations from raw pixels have been developed to replace traditional approaches which are heavily reliant on hand-engineered features. This has enabled advances in the fields of security, healthcare diagnostics, and user-interface tailoring. Innovative applications and solutions are made possible through this increased analysis of dynamic and static facial capabilities through the use of deep learning architectures such as CNNs, RNNs, LSTMs, and GANs. In this work, a comprehensive survey and review of popular deep learning techniques employed in facial character analysis are given including their main algorithmic inspirations, convolutional neural network architectures for each subcategory as well as some application specifics.

Computer vision algorithms are one of the most popular topics right now because of their importance in healthcare, industrial, and commercial applications.[4,5,6] Advancements in deep learning, notably Generative Adversarial Networks (GANs), have propelled rapid progress in facial modification studies.[7] There's a concerning trend of individuals using such technology to create deceptive or slanderous images of public figures. Various methods have emerged to categorize facial image data into different shapes. Some consider this a supervised classification task, whereas others see it as an unsupervised clustering challenge. Hairstyle and eyelash recommendations are based on a multi-model face shape as well as an eye attribute identification system.[8,9] While A Face Clustering Method Based on Facial Shape, Automatic Clustering of Faces in Meetings, and Dual threshold-based unsupervised face image clustering addressed this issue as unsupervised clustering, 3D Face data and SVM, and Face shape classification using Inception v3 addressed it as supervised classification.[10,11]

Deep-fake detection is a growing concern.[12] Researchers analyse various cues, including details in the eyes and teeth, a blend of local and global features, as well as 3D head pose, to verify face-swapped images and videos.[13] Additionally, sequential data is vital in identifying deep-fake videos.[14,15] Facial expression analysis holds promise for mental health diagnostics, as facial expressions closely reflect intentions, attitudes, and mental states.[16] Face-swapping has sparked significant discussion within the visual and graphic communities.

Because of its potential applications in a variety of fields, including film composition, privacy, computer gaming, and security, it is also important to note that improved face-swapping technology will contribute to the development of better face forgery detection tools.[17] Face and attribute recognition are utilized in a wide range of digital retail, authentication, and security applications. Face-monitoring situations typically necessitate simultaneous detection of faces, facial expressions, ages, and facial characteristics.[18] Deep network-based algorithms are one of the most important types of AI algorithms. Deep learning, a prominent topic of machine learning, uses hierarchical architectures to learn high-level and complex abstractions from data.[19] Face swapping's potential uses in a variety of fields, including computer games, movie composition, and privacy protection, have drawn a lot of interest from the vision and graphics sectors.[20]

METHOD

Data Sources

We systematically searched Scopus, Web of Science, IEEE Xplore, PubMed, and Google Scholar for relevant literature published between 2015 and 2024.

Search Strategy

Keywords used included: "Deep Learning," "Facial Recognition," "Convolutional Neural Networks," "Emotion Detection," and "Generative Adversarial Networks." Boolean operators (AND, OR) were utilized to refine the search.

Inclusion and Exclusion Criteria

Inclusion Criteria

- Studies focused on deep learning techniques for facial analysis.
- Peer-reviewed articles published in English.

Exclusion Criteria

- Non-peer-reviewed articles, editorials, and commentaries.
- Studies unrelated to facial analysis.

Deep Learning Architectures in Facial Recognition

Convolutional Neural Networks (CNNs)

CNNs: Convolutional Neural Networks (for images) Canonical for structured grid data such as Images. One of the principal reasons why Convolutional Neural Networks have become de facto standard in many computer vision tasks is their ability to automatically learn hierarchical spatial feature representations from input images. It includes the following sections with major building blocks and fundamentals of CNN.

When it comes to facial character analysis, several specific CNN architectures and models have been developed to address various aspects such as facial recognition, emotion detection, and feature extraction as shown in below mind map. Here are some notable architectures:

Figure 1. Popular CNN Architectures

Convolutional Layers:

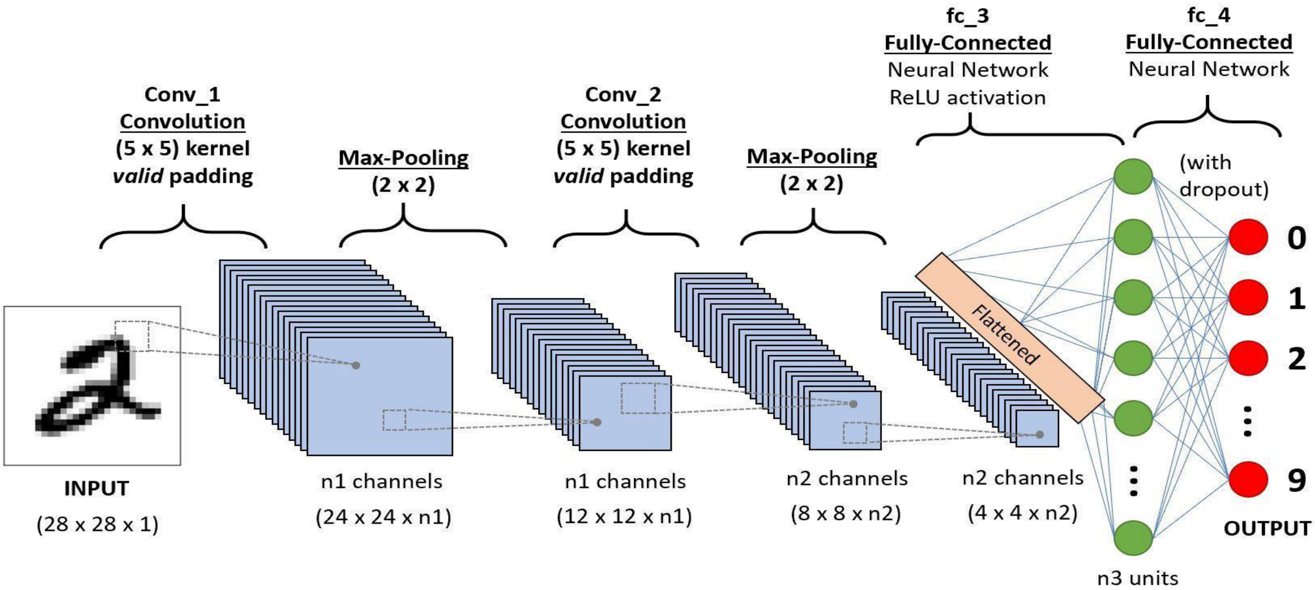
- Convolution Operation: a convolutional layer contains a set of learnable filters (or kernels) that are applied to the input image. Convolution: the input to the convolution layer is applied through a filter, so that these filters move (convolve) over an image by calculating neuron products between them with sharing weights and are of equivalent width as the receptive field in the original dimension.
- Feature Maps: on performing the convolution operation, individual feature maps are generated which depict different features like edges, textures and patterns across spatial locations.

Activation Functions:

- ReLU (Rectified Linear Unit): ReLU is an activation function that adds non-linearity to the network by zero-ing all negative values in a feature maps while keeping positive values. This helps the network to learn patterns and more complex relations between inputs.

Pooling Layers:

- Max Pooling: another type of pooling operation, it also reduces the spatial dimensions (width & height) but in this case by taking a maximum value over different windows so as to reduce computations. The window size parameter is specified when maxpool layers are created and typically max pool with 2x2 filters strides by 4 hence outputs values for each block at increments of four pixels along both width and height.) This technique is useful to reduce the number of parameters and computation along with providing translation invariance.



Source: https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/, 2020

Figure 2. Convolutional Neural Network Architecture

Output Layer: the output layer in classification tasks uses the softmax function (for multi-class classification) or sigmoid logits for binary classification purposes.

Application of CNNs to Image Analysis

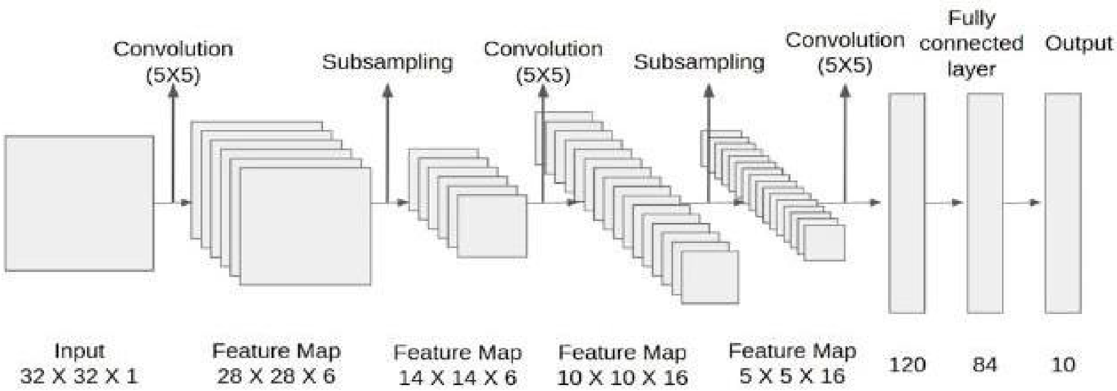
CNNs are highly effective for various image analysis tasks, including:

- Image Classification: Identifying the category of an object in an image.
- Object Detection: Detecting and localising objects within an image.
- Segmentation: Partitioning an image into segments or regions based on certain characteristics.
- Facial Recognition: Identifying and verifying individuals based on facial features.

Particular Architectures and Models for Character Analysis on the Face

When it comes to facial character analysis, several specific CNN architectures and models have been developed to address various aspects such as facial recognition, emotion detection, and feature extraction. Here are some notable architectures:

LeNet-5



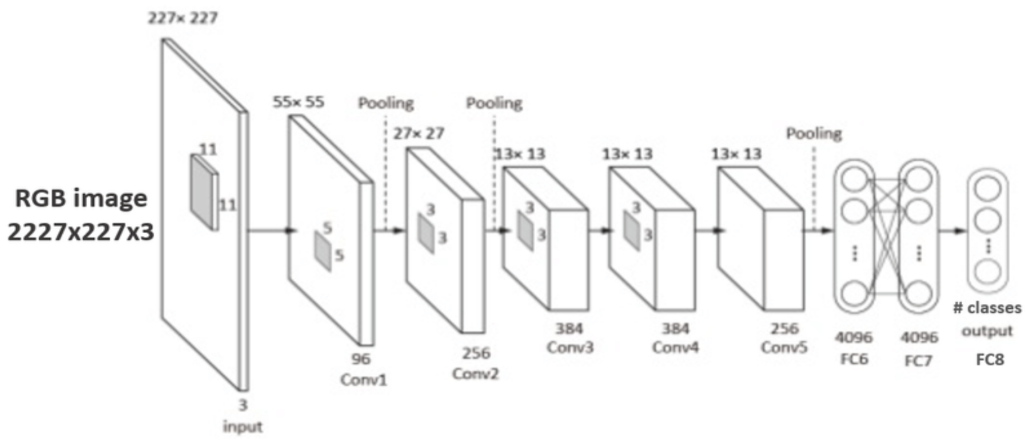
Source: https://www.analyticsvidhya.com/blog/2021/03/the-architecture-of-lenet-5/. 2021

Figure 3. LeNet Architecture

Overview: one of the oldest CNN architectures developed by Yann LeCun for digit recognition.

Application: the principles of LeNet-5 (convolutional layers followed by pooling layers) are foundational and have been further extended in more complex architectures for facial analysis.

AlexNet



Source: https://www.researchgate.net/figure/AlexNet-architecture-Includes-5-convolutional-layers-and-3-fullyconnected-layers_fig3_322592079

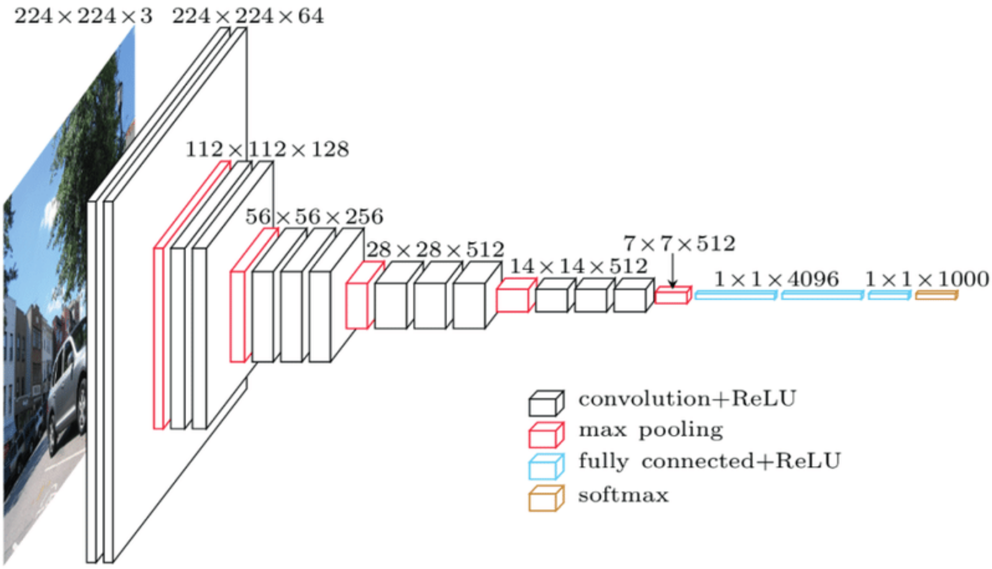
Figure 4. AlexNet Architecture

Overview: this model won the ImageNet competition in 2012 and significantly advanced the field of deep learning.

https://doi.org/10.56294/dm2025658

Application: this model demonstrated the power of deep convolutional networks in large-scale image classification tasks, such as face recognition.

VGGNet

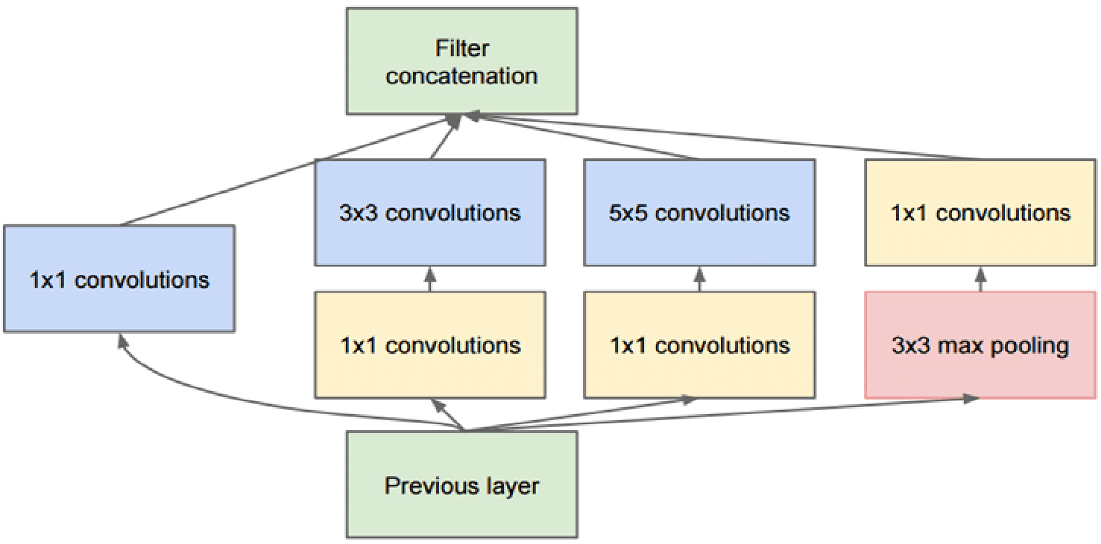


Source: https://www.researchgate.net/profile/Timea-Bezdan/publication/333242381/figure/fig2/AS:760979981860866@1558443174380/VGGNet-architecture-19.ppm
Figure 5. VGGNet Architecture

Overview: familiar for its simplicity and use of very small (3x3) convolution filters, this model architecture achieves great performance with a deeper network.

Application: generally used in facial recognition and analysis tasks due to its balanced trade-off between depth and complexity in architecture.

Inception (GoogLeNet)



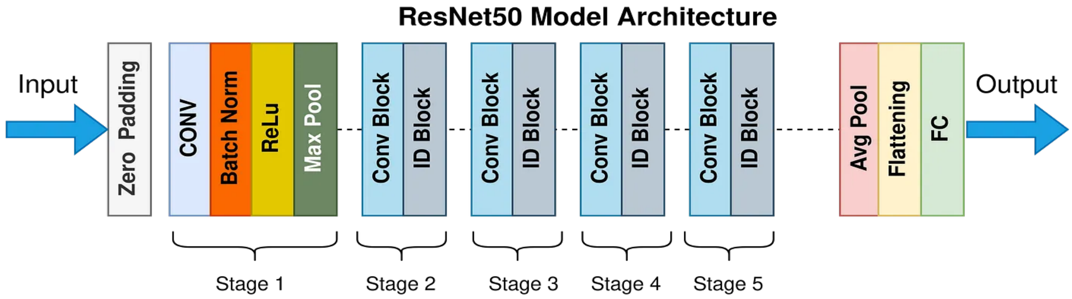
Source: https://www.researchgate.net/profile/Mahmoud-Al-Ayyoub/publication/324999549/figure/fig2/AS:624203388710912@1525833093190/GoogLeNet-Inception-Architecture.png
Figure 6. Inception (GoogLeNet) Architecture

Overview: the concept of Inception modules allow more efficient computation by varying filter sizes within the same layer.

Application: suitable for facial feature extraction and recognition, providing higher accuracy with fewer parameters compared to VGGNet model.

https://doi.org/10.56294/dm2025658

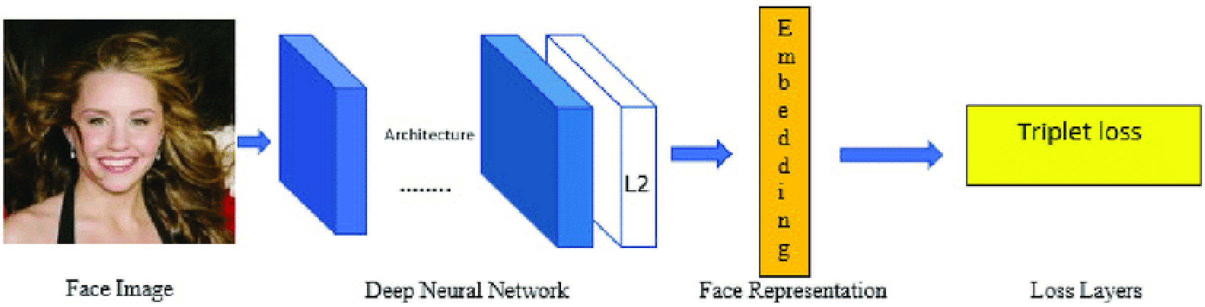
ResNet (Residual Networks)



Source: https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758
Figure 7. ResNet50 Architecture

Overview: this ResNet model is introduced by Microsoft, which addresses the vanishing gradient problem by skipping the connections (residual connections) to allow gradients to flow more easily through deeper networks.
Application: popularly used in facial recognition systems, employed by security agencies and large-scale tech companies.

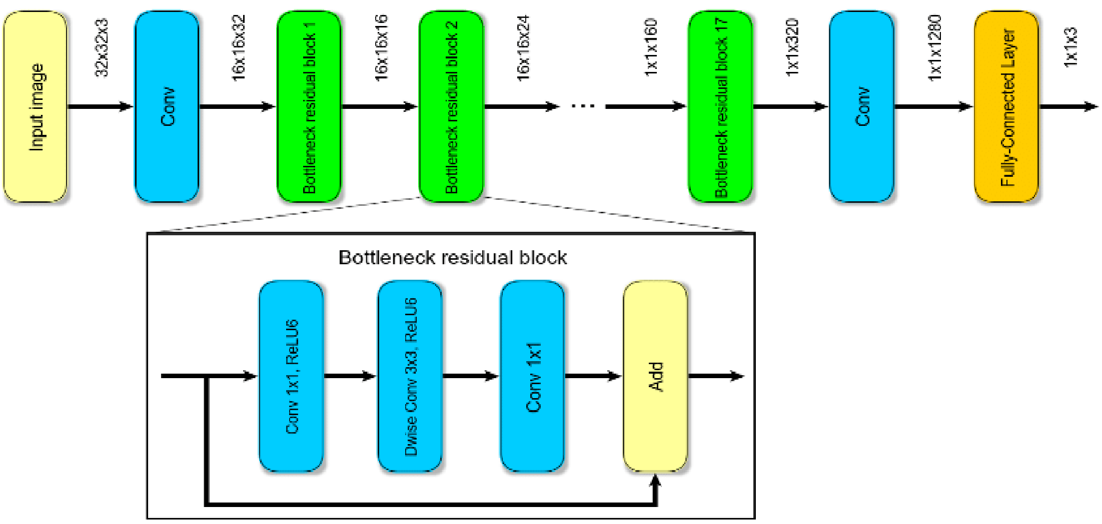
FaceNet



Source: https://www.researchgate.net/publication/362474429/figure/fig2/AS:11431281081022709@1661464596934/FaceNet-model-architecture-FaceNet-consists-of-a-batch-input-layer-and-a-deep-CNN-DCNN.png
Figure 8. FaceNet Architecture

Overview: introduced by Google, FaceNet uses a deep convolutional network to directly learn a mapping from face images to a compact Euclidean space to measure face similarity based on distances between facial features.
Application: used for face recognition, verification, and clustering tasks with high precision.

MobileNet

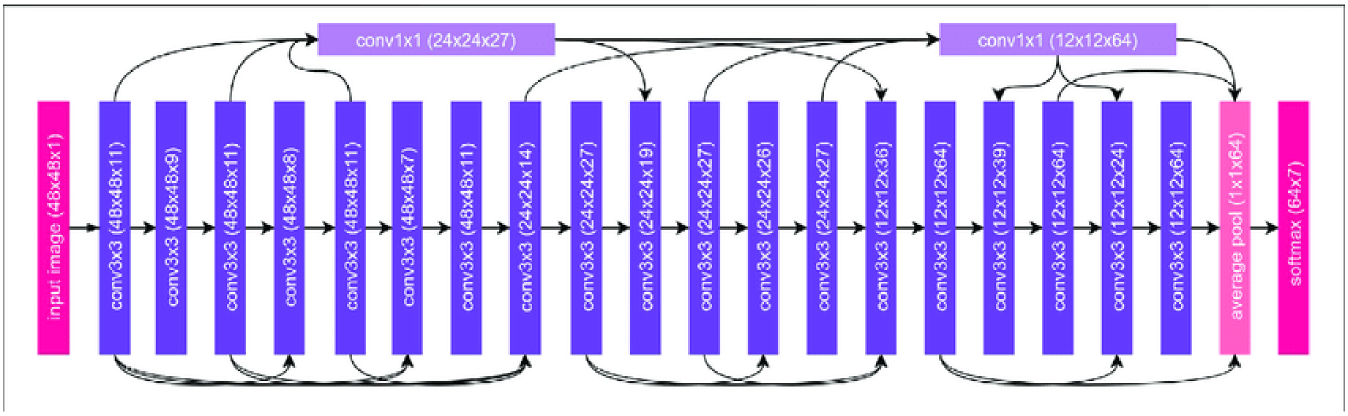


Source: https://www.researchgate.net/figure/The-architecture-of-the-MobileNetv2-network_fig3_342856036
Figure 9. MobileNet Architecture

Overview: this model is designed for efficient computation, MobileNet is optimized for embedded and mobile vision applications.

Application: ideal for real-time facial analysis on equipment with limited computational power, such as embedded systems and smartphones.

Emotion Net



Source: https://www.researchgate.net/publication/369511271/figure/fig2/AS:11431281130138885@1679753444018/Schematic-Diagram-of-Emotion-Recognition-Module.png

Figure 10. Emotion Net Architecture

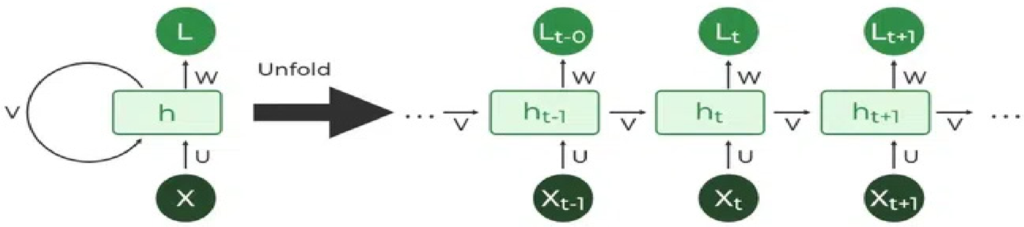
Overview: this model architecture is focused on recognizing facial emotions and expressions.

Application: this model is used in applications such as mental health monitoring to enhance user experience in human-computer interaction.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTMs)

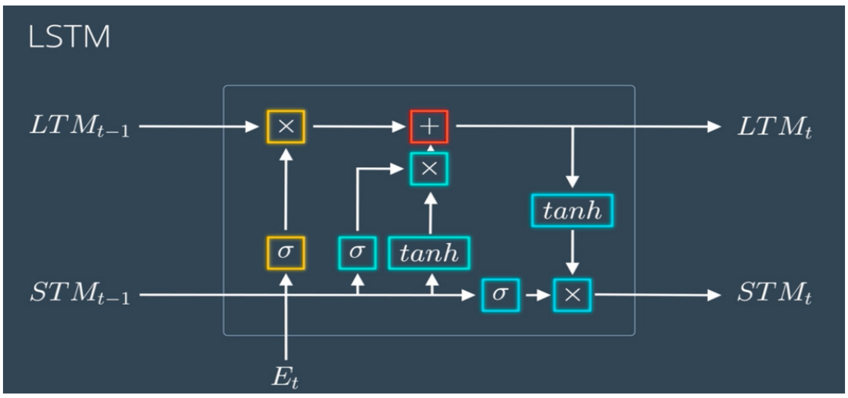
Basics of RNNs and LSTMs

Recurrent Neural Networks (RNNs) are a class of neural networks designed to handle sequential data by maintaining a 'memory' of previous inputs in their internal state. This makes RNNs particularly suited for tasks involving time-series data, where the order of the data points is crucial. However, standard RNNs struggle with long-term dependencies due to issues like vanishing and exploding gradients.



Source: https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/

Figure 11. Recurrent Neural Networks Architecture



Source: https://www.analyticsvidhya.com/blog/2021/01/understanding-architecture-of-lstm/, 2021

Figure 12. Long Short-Term Memory Networks Architecture

https://doi.org/10.56294/dm2025658

To address these limitations, Long Short-Term Memory Networks (LSTMs) were developed. LSTMs are a special type of RNN capable of learning long-term dependencies. They achieve this by incorporating memory cells and gating mechanisms that control the flow of information.

Key Components of LSTMs

1. Cell State: the cell state acts as a conveyor belt, carrying relevant information across many time steps. It's regulated by gates to retain or discard information.
2. Gates: LSTMs use three gates to manage information:
   - Forget Gate: decides what information to discard from the cell state.
   - Input Gate: determines which new information to add to the cell state.
   - Output Gate: controls the output based on the cell state.

Application of RNNs and LSTMs for Recording Facial Expressions' Temporal Patterns and Sequence

Facial expressions are dynamic and change over time. To accurately analyse and interpret these expressions, it is essential to capture their temporal patterns and sequences. RNNs and LSTMs are well-suited for this task due to their ability to process sequential data and remember previous states. Here's how they can be applied to facial expression analysis:

1. Temporal Dynamics of Facial Expressions:
   - Facial expressions involve subtle and rapid changes in facial muscles over time. RNNs and LSTMs can capture these temporal dynamics, enabling the detection of nuanced expressions and transitions between different emotional states.

2. Sequence Modelling:
   - RNNs and LSTMs can model sequences of facial expressions, allowing for the recognition of complex emotion patterns that unfold over multiple frames in a video. This is crucial for understanding emotions like surprise, which may involve a rapid succession of expressions.

3. Emotion Recognition:
   - By processing sequences of facial images, LSTMs can recognize emotions based on how facial features evolve over time. This improves the accuracy of emotion detection systems, as emotions are often expressed through a series of facial movements rather than static snapshots.

4. Facial Action Unit (FAU) Detection:
   - FAUs are specific facial muscle movements associated with expressions. RNNs and LSTMs can track these movements across frames, providing a detailed analysis of how different parts of the face contribute to an expression.

Specific Models and Techniques

Several models and techniques utilise RNNs and LSTMs for facial expression analysis:

1. Deep Temporal Models:
   - Deep Temporal Convolutional Networks: Combine convolutional layers to extract spatial features and LSTM layers to capture temporal dependencies. These models are effective for analysing facial expressions in video sequences.

2. Attention Mechanisms:
   - Attention mechanisms can be incorporated into LSTM networks to focus on specific parts of the sequence that are most relevant for emotion recognition. This helps in dealing with longer sequences by prioritising important frames.

3. Hybrid Models:
   - CNN-LSTM Models: Combine Convolutional Neural Networks (CNNs) with LSTMs to leverage both spatial and temporal information. CNNs extract spatial features from each frame, while LSTMs model the temporal relationships between frames.

Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs) are a class of machine learning frameworks introduced by Ian Goodfellow and his colleagues in 2014. GANs consist of two neural networks, the generator and the discriminator, that are trained simultaneously through a process of adversarial training:

1. Generator (G):

- Objective: to generate synthetic data that is indistinguishable from real data.
- Mechanism: takes random noise as input and generates synthetic data.

2. Discriminator (D):
   - Objective: to distinguish between real data and the synthetic data generated by the generator.
   - Mechanism: receives both real and synthetic data as input and outputs a probability indicating whether the input data is real or fake.

The training process involves the generator trying to fool the discriminator by producing realistic data, while the discriminator tries to improve its ability to identify real versus fake data. This adversarial process continues until the generator produces highly realistic synthetic data that the discriminator can no longer distinguish from real data.

Role of GANs in Generating Synthetic Facial Data

GANs play a significant role in generating synthetic facial data, which is crucial for various applications, including data augmentation, privacy preservation, and enhancing analysis. Here are some key contributions:

1. Data Augmentation:
   - Enhancing Training Datasets: GANs can generate large amounts of synthetic facial data to augment training datasets. This is particularly useful when the available data is limited, diverse, or imbalanced.
   - Variations and Realism: GANs can create diverse facial images with variations in expressions, poses, lighting, and occlusions, enhancing the robustness of models trained on these datasets.

2. Privacy Preservation:
   - Anonymization: synthetic facial data generated by GANs can be used to create anonymized datasets that retain the statistical properties of the original data but do not reveal the identities of real individuals.
   - Synthetic Identities: GANs can generate entirely new synthetic identities for testing and validation purposes, protecting the privacy of real individuals.

3. Data Imputation and Enhancement:
   - Missing Data: GANs can be used to impute missing facial data in images, such as reconstructing occluded parts of the face.
   - Image Enhancement: GANs can improve the quality of facial images by removing noise, increasing resolution, and correcting distortions.

Role of GANs in Enhancing Analysis

Beyond data generation, GANs also enhance the analysis of facial data in various ways:

1. Facial Expression Synthesis:
   - Emotion Manipulation: GANs can be used to manipulate facial expressions, generating images that depict different emotions. This capability can help in training and evaluating emotion recognition systems.
   - Dynamic Expressions: GANs can generate sequences of facial images that simulate dynamic expressions, providing rich datasets for studying temporal patterns in facial expressions.

2. Feature Extraction and Representation Learning:
   - Unsupervised Learning: GANs can learn meaningful representations of facial features without requiring labeled data, which can then be used for various downstream tasks like recognition and verification.
   - Latent Space Exploration: By exploring the latent space of GANs, researchers can understand and visualize how different facial attributes (e.g., age, gender, expression) are encoded, aiding in interpretability and analysis.

3. Adversarial Training and Robustness:
   - Improving Model Robustness: GANs can generate adversarial examples to test the robustness of facial recognition and analysis systems. This helps in identifying and mitigating vulnerabilities in the models.
   - Defensive Mechanisms: GANs can be used to develop and test defensive mechanisms against adversarial attacks, ensuring the reliability and security of facial analysis systems.

4. Style Transfer and Image Translation:
    - Cross-Domain Analysis: GANs can perform tasks like facial style transfer and domain adaptation, enabling the analysis of facial images across different domains (e.g., sketches to photos, day to night).
    - Attribute Manipulation: GANs can modify specific facial attributes (e.g., adding glasses, changing hair color) in images, which is useful for studying the impact of these attributes on facial recognition systems.

Hybrid Models

Hybrid models in deep learning combine different types of neural network architectures and strategies to leverage their complementary strengths and mitigate their individual limitations. This approach often results in models that are more accurate, robust, and efficient. Hybrid models can be particularly beneficial for complex tasks such as facial character analysis, where capturing both spatial and temporal features is crucial.

Components of Hybrid Models

Hybrid models often integrate elements from various neural network architectures, including:

1. Convolutional Neural Networks (CNNs):
    - Strengths: excellent at capturing spatial hierarchies in data, making them ideal for tasks involving image analysis.
    - Applications: feature extraction from facial images, recognizing patterns and details.

2. Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTMs):
    - Strengths: capable of modelling sequential data and capturing temporal dependencies.
    - Applications: analyzing sequences of facial expressions over time, understanding dynamic changes.

3. Generative Adversarial Networks (GANs):
    - Strengths: effective in generating synthetic data, improving the diversity and size of training datasets.
    - Applications: data augmentation, enhancing training data quality.

4. Attention Mechanisms:
    - Strengths: improve the model's focus on relevant parts of the input, enhancing interpretability and performance.
    - Applications: facial feature localization, emotion recognition.

5. Autoencoders:
    - Strengths: useful for unsupervised learning and data compression.
    - Applications: feature learning, data denoising.

Hybrid Models for Facial Character Analysis

Combining the strengths of these different models can significantly enhance the accuracy and robustness of facial character analysis systems.

Figure 13. Popular Hybrid Models for Facial Character Analysis

Here are some popular hybrid approaches:

1. CNN-RNN/LSTM Models:
    - Architecture: convolutional layers for spatial feature extraction followed by recurrent layers for temporal pattern recognition.
    - Application: emotion detection in video sequences, where facial features are extracted by CNNs and the temporal evolution of these features is modelled by RNNs or LSTMs.
    - Example: using CNNs to process each frame of a video to extract facial features, which are then fed into an LSTM to capture the temporal dynamics of expressions.

2. CNN-GAN Models:

https://doi.org/10.56294/dm2025658

- Architecture: CNNs for feature extraction combined with GANs for generating synthetic facial data.
- Application: data augmentation for training deep learning models, improving the diversity and realism of the training dataset.
- Example: training a GAN to generate synthetic facial expressions that are used to augment the dataset for a CNN-based emotion recognition model.

3. Attention-CNN Models:
   - Architecture: incorporating attention mechanisms within CNN architectures to enhance the model's focus on critical regions of the face.
   - Application: improving the detection of subtle facial expressions and features by directing the model's attention to important areas.
   - Example: using attention layers to weigh different parts of the facial image more heavily during processing, improving accuracy in recognizing emotions or identifying individuals.

4. Autoencoder-RNN Models:
   - Architecture: using autoencoders for unsupervised pre-training followed by RNNs or LSTMs for sequence prediction.
   - Application: feature learning and sequence modelling in tasks where labelled data is scarce.
   - Example: training an autoencoder on unlabeled facial images to learn a compact representation, then using these representations as input to an LSTM for emotion prediction in video sequences.

5. Multimodal Hybrid Models:
   - Architecture: combining different types of data (e.g., visual, auditory) and processing them with specialized networks before merging the information.
   - Application: enhancing facial character analysis by incorporating additional data modalities.
   - Example: integrating facial image analysis (via CNNs) with audio emotion analysis (via RNNs) to improve overall emotion recognition accuracy.

Benefits of Hybrid Models: Improved Accuracy, Enhanced Robustness, Increased Efficiency and Greater Flexibility.

Transfer Learning

Transfer learning is a machine learning technique where a model developed for a particular task is reused as the starting point for a model on a different but related task. This approach leverages pre-trained models, which have been trained on large datasets, to improve the performance and efficiency of models for specific tasks with limited data. Transfer learning is especially useful in deep learning due to the high computational cost and extensive data requirements of training deep neural networks from scratch.

Key Concepts of Transfer Learning

1. Pre-trained Models:
   - Definition: models that have been previously trained on large datasets, such as ImageNet, which contains millions of labelled images.
   - Examples: popular pre-trained models include VGGNet, ResNet, Inception, and BERT for natural language processing.

2. Feature Extraction:
   - Process: using the convolutional layers of a pre-trained model to extract features from new data. The extracted features are then fed into a new classifier (typically fully connected layers) that is trained on the specific task.

3. Fine-tuning:
   - Process: starting with a pre-trained model and fine-tuning its parameters on the new task's data. This involves unfreezing some of the top layers of the base model and jointly training both the newly added classifier layers and the base model layers.

4. Domain Adaptation:
   - Definition: adjusting the model to work well on a new domain that may have different characteristics from the domain it was originally trained on.

https://doi.org/10.56294/dm2025658

Application of Transfer Learning in Facial Character Analysis

Transfer learning is highly effective for facial character analysis, where obtaining a large, labelled dataset can be challenging. Here's how transfer learning can be applied and modified for this purpose:

1. Using Pre-trained CNNs for Feature Extraction.
2. Fine-tuning Pre-trained Models.
3. Combining Transfer Learning with RNNs for Temporal Analysis.
4. Domain Adaptation Techniques.
5. Transfer Learning for Specific Attributes.

Benefits of Transfer Learning: Reduced Training Time, Improved Performance, Generalization, Effective Use of Small Datasets and Versatility Across Domains.

Supplementary Methods

As deep learning continues to evolve, several emerging techniques are gaining attention for their potential to address complex problems and improve the accuracy and efficiency of models. While these methods may not yet be as widely adopted as traditional approaches, they hold significant promise for advancing the field. Here, we provide an overview of some of these emerging deep learning techniques as shown in below mind map.

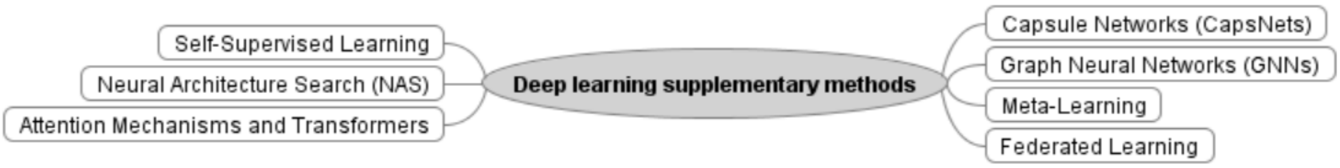


Figure 14. Popular Deep Learning Supplementary Methods

Capsule Networks (CapsNets)

Overview: Capsule Networks, introduced by Geoffrey Hinton and his team, aim to address some limitations of Convolutional Neural Networks (CNNs), particularly their inability to capture spatial hierarchies between features.

Key Features:

- Capsules: groups of neurons that represent different properties of objects or parts of objects, such as position, orientation, and scale.
- Dynamic Routing: mechanism to ensure that capsules at one level send their outputs to appropriate capsules at the next level.

Applications:

- Improved image recognition, especially for tasks requiring detailed spatial relationships.
- Robustness to adversarial attacks and variations in orientation and perspective.

Graph Neural Networks (GNNs)

Overview: Graph Neural Networks are designed to handle graph-structured data, which consists of nodes (entities) and edges (relationships). They are particularly useful for tasks where data is inherently non-Euclidean.

Key Features:

- Node Embeddings: representing each node by aggregating information from its neighbors.
- Graph Convolution: generalizing convolution operations to graph data, enabling the learning of node representations based on the graph structure.

Applications:

- Social network analysis, recommendation systems, and protein interaction networks.
- Semi-supervised learning on graph data, such as citation networks and knowledge graphs.

Self-Supervised Learning

Overview: self-supervised learning leverages unlabeled data by creating surrogate tasks where the supervision signal is derived from the data itself. This approach reduces the reliance on large labelled datasets.

Key Features:

- Pretext Tasks: tasks like predicting rotations, filling in missing parts of an image, or predicting future frames in a video.
- Representation Learning: learning useful features from large amounts of unlabeled data, which can be fine-tuned on downstream tasks.

https://doi.org/10.56294/dm2025658

Applications:

- Natural language processing (e.g., BERT, GPT) and computer vision (e.g., contrastive learning).
- Enhancing performance in scenarios with limited labelled data.

Neural Architecture Search (NAS)

Overview: Neural Architecture Search automates the design of neural network architectures, aiming to find optimal models for specific tasks without manual intervention.

Key Features:

- Search Space: defines possible network architectures, including various layers, connections, and hyperparameters.
- Search Strategy: uses techniques like reinforcement learning, evolutionary algorithms, or gradient-based methods to explore the search space.
- Performance Estimation: evaluates the performance of candidate architectures using proxy metrics or partial training.

Applications:

- Image classification, object detection, and language modelling.
- Customizing architectures for hardware constraints, such as mobile devices.

Meta-Learning

Overview: meta-learning, or "learning to learn," focuses on developing models that can quickly adapt to new tasks with minimal data. It aims to improve the generalization capabilities of learning algorithms.

Key Features:

- Few-Shot Learning: enabling models to learn from a few examples by leveraging prior knowledge.
- Model-Agnostic Meta-Learning (MAML): an algorithm that optimizes models to be easily adaptable to new tasks.

Applications:

- Personalized recommendations, adaptive control systems, and medical diagnosis.
- Rapid adaptation to new environments and tasks in robotics and autonomous systems.

Attention Mechanisms and Transformers

Overview: attention mechanisms have revolutionized many areas of deep learning by allowing models to focus on relevant parts of the input data. Transformers, which rely heavily on attention mechanisms, have become the backbone of state-of-the-art models in natural language processing.

Key Features:

- Self-Attention: enables the model to weigh the importance of different parts of the input sequence, capturing long-range dependencies.
- Transformers: composed of multiple layers of self-attention and feedforward networks, enabling parallel processing and scalability.

Applications:

- Natural language understanding and generation (e.g., BERT, GPT, T5).
- Image processing tasks like image generation and object detection.

Federated Learning

Overview: Federated Learning is a decentralized approach where models are trained across multiple devices or servers holding local data samples, without exchanging the data itself.

Key Features:

- Privacy-Preserving: enhances privacy by keeping data localized and only sharing model updates.
- Scalability: capable of leveraging vast amounts of decentralized data from devices like smartphones and IoT devices.

Applications:

- Collaborative learning scenarios in healthcare, finance, and mobile applications.
- Training models on sensitive data without compromising privacy.

Recent Advances and Trends

Increasing Accuracy and Efficiency

The fields of machine learning and deep learning have seen remarkable improvements in both accuracy and

https://doi.org/10.56294/dm2025658

efficiency, driven by advances in algorithms and computational power. These improvements have enabled the development of more sophisticated models and applications across various industries.

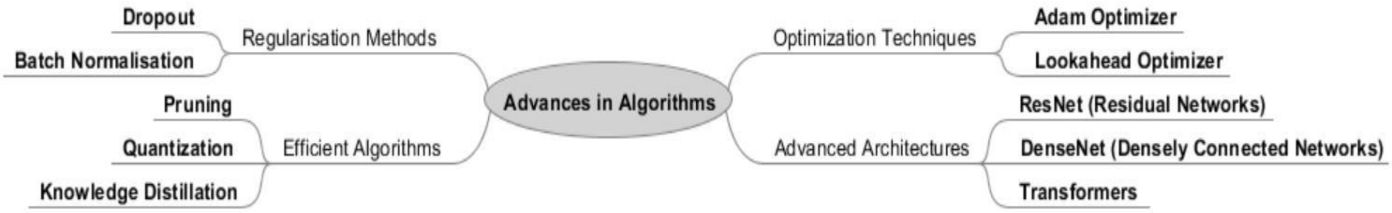


Figure 15. Advancements in Algorithms in Machine Learning and Deep Learning

Advances in Algorithms

1. Optimization Techniques:
   - Adam Optimizer: combines the advantages of AdaGrad and RMSProp, providing an adaptive learning rate that enhances convergence speed and accuracy.
   - Lookahead Optimizer: enhances optimization by looking ahead at the potential steps and guiding the search direction, leading to faster and more stable convergence.

2. Regularization Methods:
   - Dropout: randomly drops units from the neural network during training, preventing overfitting and improving generalization.
   - Batch Normalization: normalize the inputs of each layer to stabilize learning and reduce the training time, leading to improved accuracy.

3. Advanced Architectures:
   - ResNet (Residual Networks): Introduces skip connections to alleviate the vanishing gradient problem, enabling the training of very deep networks.
   - DenseNet (Densely Connected Networks): each layer receives input from all previous layers, promoting feature reuse and reducing the number of parameters.
   - Transformers: use self-attention mechanisms to handle dependencies and parallelize training, leading to breakthroughs in natural language processing and other sequential tasks.

4. Efficient Algorithms:
   - Pruning: reduces the size of neural networks by removing less important connections, leading to faster inference without significant loss in accuracy.
   - Quantization: reduces the precision of the weights and activations, allowing models to run efficiently on hardware with limited computational resources.
   - Knowledge Distillation: trains a smaller, efficient model (student) using the output of a larger, pre-trained model (teacher), maintaining high performance with fewer resources.

Advances in Computational Power

1. Hardware Accelerators:
   - GPUs (Graphics Processing Units): provide massive parallel processing capabilities, significantly accelerating the training and inference of deep learning models.
   - TPUs (Tensor Processing Units): custom accelerators developed by Google specifically for machine learning workloads, offering high performance and efficiency.
   - FPGAs (Field-Programmable Gate Arrays): offer customizable hardware acceleration, providing a balance between performance and flexibility.

2. Distributed Computing:
   - Cluster Computing: uses multiple interconnected computers to distribute the workload, allowing the training of large models on massive datasets.
   - Cloud Computing: provides scalable resources on-demand, enabling researchers and practitioners to access high-performance computing without significant upfront investment.

3. Efficient Software Frameworks:
   - TensorFlow and PyTorch: popular deep learning frameworks that provide optimized libraries, tools, and APIs for building and deploying models efficiently.

https://doi.org/10.56294/dm2025658

- ONNX (Open Neural Network Exchange): an open format for representing machine learning models, enabling interoperability between different frameworks and optimizing deployment.

4. Parallelization Techniques:
   - Data Parallelism: distributes data across multiple processors, allowing simultaneous training of different batches of data.
   - Model Parallelism: distributes different parts of a model across multiple processors, enabling the training of very large models that cannot fit into the memory of a single device.

Combining Algorithmic and Computational Advances

The combination of algorithmic innovations and enhanced computational power has led to significant improvements in the accuracy and efficiency of deep learning models. Some notable examples include:

- GPT-3: a large language model with 175 billion parameters, trained using advanced optimization techniques and distributed computing resources, showcasing state-of-the-art performance in natural language processing tasks.
- Efficient Net: a family of models that use a compound scaling method to achieve state-of-the- art accuracy on image classification tasks with fewer parameters and lower computational cost compared to previous models.
- AlphaFold: uses deep learning to predict protein structures with high accuracy, leveraging advanced architectures, efficient algorithms, and powerful hardware accelerators.

Real-time Analysis

Real-time analysis of facial data has numerous applications, including security and surveillance, human-computer interaction, virtual reality, and healthcare. The development of systems capable of processing facial data in real-time requires advances in algorithms, hardware, and software to ensure high-speed performance without compromising accuracy.

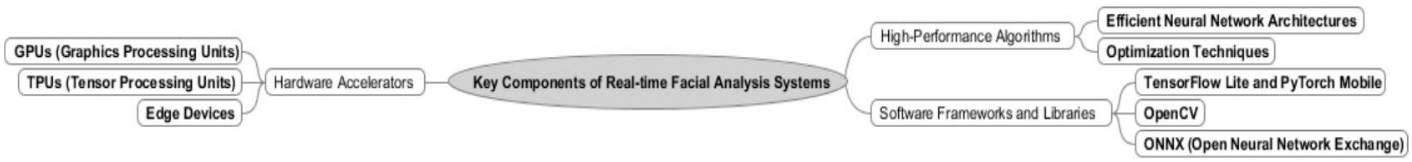


Figure 16. Key Components of Real-time Facial Analysis Systems

Steps in Developing Real-time Facial Analysis Systems

1. Data Acquisition and Preprocessing:
   - High-Quality Datasets: collecting diverse and representative datasets to train robust models.
   - Preprocessing Techniques: techniques such as face alignment, normalisation, and augmentation to enhance the quality and variability of the training data.

2. Model Training and Optimization:
   - Training Efficient Models: using architectures designed for real-time applications, such as MobileNet, YOLO (You Only Look Once), and SSD (Single Shot MultiBox Detector).
   - Transfer Learning: leveraging pre-trained models and fine-tuning them on specific tasks to reduce training time and improve performance.
   - Model Compression: applying techniques like pruning, quantization, and knowledge distillation to reduce the size and complexity of models, making them suitable for real- time deployment.

3. Deployment on Hardware:
   - Edge Devices: deploying models on edge devices to ensure low latency and high- speed processing. Edge devices like NVIDIA Jetson Nano, Google Coral, and Intel Movidius are commonly used for this purpose.
   - Cloud-based Solutions: utilizing cloud services for scenarios where edge deployment is not feasible. Cloud providers offer powerful infrastructure for real-time processing but may introduce latency due to data transmission.

4. Post-processing and Output:
   - Real-time Visualization: displaying the results of facial analysis in real-time, such as emotion recognition, face identification, and expression tracking.

- Actionable Insights: generating actionable insights based on the analysis, such as triggering alerts in security systems or adapting interactions in human-computer interfaces.

Applications of Real-time Facial Analysis

Figure 17. Various Applications of Real-time Facial Analysis

1. Security and Surveillance:
   - Real-time Face Recognition: Identifying individuals in real-time to enhance security measures in public and private spaces.
   - Anomaly Detection: Monitoring facial expressions and behaviors to detect suspicious activities.

2. Human-Computer Interaction:
   - Emotion Recognition: analyzing facial expressions to adapt the behavior of digital assistants, virtual agents, and interactive systems.
   - User Authentication: Implementing real-time facial recognition for secure and convenient user authentication.

3. Virtual Reality and Augmented Reality:
   - Facial Tracking: Enabling realistic avatar interactions by capturing and replicating facial expressions in virtual environments.
   - Augmented Experiences: Enhancing AR applications with real-time facial analysis, such as virtual try-ons and interactive advertising.

4. Healthcare:
   - Patient Monitoring: Real-time analysis of facial expressions to monitor pain levels, emotions, and other health indicators.
   - Telemedicine: improving remote consultations by providing real-time insights into patients' emotional and physical states.

Multi-Modal Systems

Multi-modal biometric systems combine multiple types of biometric data to improve the accuracy, robustness, and security of identification and authentication processes. Integrating facial analysis with other biometric modalities, such as voice, fingerprint, iris, and gait, enhances the system's overall performance by leveraging the strengths of each modality.

Benefits of Multi-Modal Systems: Improved Accuracy, Enhanced Robustness and Increased Security.

Figure 18. Applications of Multi-Modal Systems

Key Components of Multi-Modal Systems

Data Acquisition:

- Facial Analysis: capturing facial images or video streams using cameras.
- Voice Recognition: recording audio using microphones to analyze speech patterns.
- Fingerprint Scanning: using fingerprint sensors to capture ridge patterns.
- Iris Recognition: capturing high-resolution images of the iris.

- Gait Analysis: analyzing body movement patterns through video or sensors.

Feature Extraction:

- Facial Features: extracting key points, textures, and shapes from facial images.
- Voice Features: analyzing pitch, tone, and speech rhythm.
- Fingerprint Features: identifying minutiae points such as ridges and bifurcations.
- Iris Features: capturing the unique patterns of the iris.
- Gait Features: analyzing stride length, walking speed, and limb movement patterns.

Feature Fusion:

- Early Fusion: combining raw data or features at the initial stage before processing.
- Late Fusion: integrating the results of individual biometric analyses at the decision level.
- Hybrid Fusion: combining both early and late fusion techniques to leverage the advantages of each approach.

Classification and Matching:

- Machine Learning Models: training models to classify and match combined biometric data.
- Similarity Measures: calculating similarity scores for different modalities and combining them to make final decisions.

RESULTS

The use of several deep learning models has changed how we recognize and analyse a face. CNNs are still the main backbone of many networks, because of their strong feature extraction powers. While CNNs focus on spatial aspects of facial expressions (as needed for emotion detection), RNNs and LSTMs capture temporal dynamics of facial expression changes. In addition to improving datasets, GANs assist with data generation that preserves the privacy of users. They combine multiple architectures and take advantage of their strengths to create hybrid models to tackle the complexity of the task that needs to be done. Nevertheless, issues remain with tackling data bias, achieving model fairness, and fulfilling real-time processing requirements. Stringent safeguards are essential to address the ethical and legal issues surrounding data privacy and misuse.

Existing research on face shape detection has various merits and demerits, including computational risk, a focus on front-angle faces, longer calculation times, fewer parameters, and the use of small datasets. Several recent methods for classifying face shape detection are analysed and given here on table 1.

One of the latest research papers on face shape detection with paper title “MML-FuseNet: An Efficient Multi-Task Mutual Learning Framework for Face Shape Detection Based on Confined Feature Fusion” outcome is shared here. This research paperwork is under final review with Computational Intelligence journal for publication.

In this study, Swin-PCT-LSTM is utilized for classification to improve the system’s effectiveness and accuracy. This categorization technique generates a useful system for categorizing different face shapes. This categorization is achieved through a series of processes, including pre-processing, identifying face landmarks and boundaries, feature extraction, and weight assignment.

The Face Shape dataset measures the proposed model parameter accuracy at 98,96 %, precision at 97,41 %, recall at 97,4 %, F1 score at 97,53 %, RSME at 0,035 %, MAE at 0,05 %, FPR at 0,0125 %, FNR at 0,03 %, TNR at 0,97, PCC at 0,94 %, Dice score at 99,42, ROC at 0,99 %, and Jaccard coefficient at 98 % respectively.

The model’s classification outcome is described in table 2 for women’s Face shape and figure 19 for Men’s face shape dataset. As a result, the different shapes may be classified more precisely with less time.

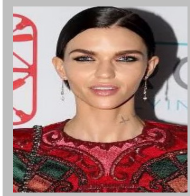
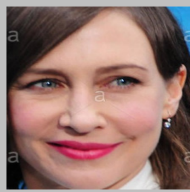
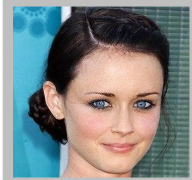
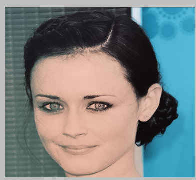
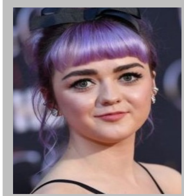
Table 1. Analysis of existing methods with its performance

No.	Author Name and references	Year of publication	Implementation Techniques	Research Objective	Model Limitation	Model Performance
1	Sarakon et al.[40]	2014	3D face data and SVM	Face shape classification from 3D human data by using SVM	Limited dataset size and Computational complexity	Accuracy 73,68 %
2	Setyadi et al.	2015	ANN	The extracted features and personality of the human can be detected using an artificial neural network (ANN).	This method highly suffered due to over fitting issues and requires larger datasets	Accuracy 42,5 %
3	Zafar et al.[39]	2016	The geometric features are used for eye-ware and face classification for data-driven eye-ware recommendation systems.	In this method, extraction algorithm eye-ware shapes are used to extract the shape of the polygonal eye exactly.	This method suffered due to a lack of detection of other facial features in the human body.	Accuracy 90 %
4	Bansode et al.[41]	2016	Region Similarity, Correlation and Fractal Dimensions	Face shape classification based on region similarity, correlation and fractal dimensions	Sensitivity to Lighting and Pose Variations. & Limited dataset diversity.	Accuracy 80 %
5	Sunhem et al.[42]	2016	Active Appearance Model (AAM), segmentation, and SVM	An Approach to Face Shape Classification for Hairstyle Recommendation	Limited Hairstyle Database & Influence of External Factors like head pose, facial expressions, etc	Accuracy 72 %
6	Rahmat et al.[43]	2018	Probability Neural Network (PNN) and Invariant Moments	Probabilistic Neural Network and Invariant Moments for Men Face Shape Classification	Gender-Specific Focus & Sensitivity to Image Quality	Accuracy 80 %
7	Petpairote et al.	2018	Personalized face neutralization using the best- matched face shape with the neutral-face database.	For wrapping the textures and neutralizing the face expression, the best-matched face-shape template was utilized	This method suffered due to low accuracy and noisy outcomes.	Accuracy 95,06 %
8	Tio et al.[44]	2019	Inception V3 Deep Learning Model	To develop a high-accuracy face shape classification system using the Inception V3 deep learning model.	High Computational Requirements & Overfitting Risk	Accuracy 84,4 %
9	Pasupa et al.[45]	2019	Hybrid approach VGG and SVM	To develop an effective face shape classification system to enhance hairstyle recommendation services.	Complexity of Hybrid Models; Dependence on Dataset Quality; Limited Hairstyle Options	Accuracy 70,3 %
10	VenkateswarLal et al.[28]	2019	SVM	To focus on texture and color features for effective face recognition	Overfitting issue	Accuracies of 99 % and 94 % for FERET data samples and Labeled Faces in the Wild data samples
11	Zhao et al.[24]	2020	ML	To predict facial structure, skin texture, and face shape structural features	A small amount of dataset.	91,0 % of accuracy

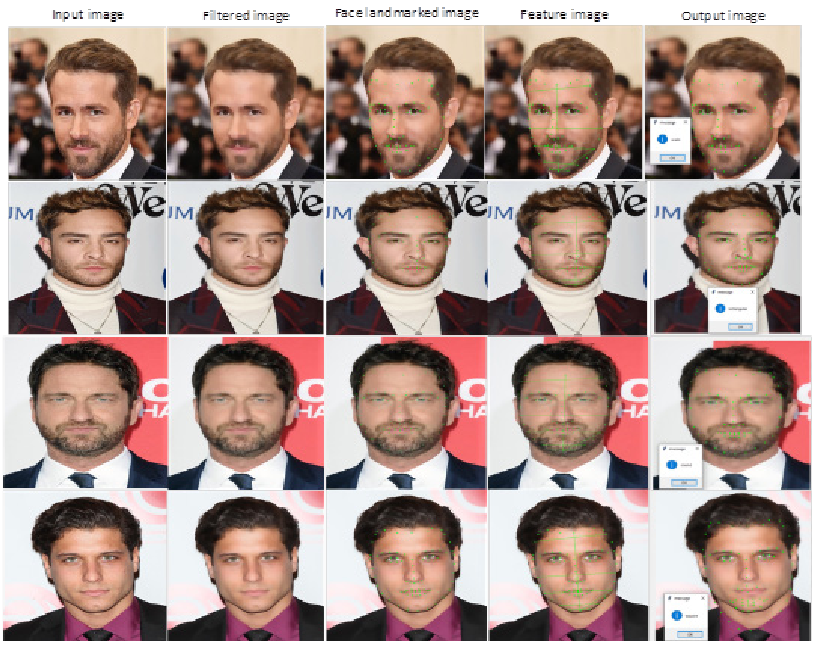
https://doi.org/10.56294/dm2025658

12	Wulansari et al.	2021	Based on the canny method, someone's character is detected.	The face shape such as long diamond, square, oval, round and heart are detected using this canny method	This method suffered due to poor accuracy.	80,0 % of accuracy
13	Shao et al.[25]	2021	JAA-Net	To detect the facial alignment effectively.	Computational risk.	94,0 % of accuracy.
14	Sukumaran et al.[22]	2021	FS-GU	To detect the face shape by considering relevant features.	Consumes more time.	94,3 % of accuracy.
15	Alzahrani et al. [23]	2021	Hybrid framework of handcrafting and learned features.	To determine eye characteristics, categorize facial forms, or determine gender.	Slowest convergence and computational complexity.	85,6 % of accuracy
16	Mehta et al.[21]	2022	ML	To classify the face shape effectively.	Only classify the front angle faces.	70 % of accuracy
17	Adityatama et al.[26]	2023	Convolutional Neural Network Xception Architecture with Transfer Learning.	To classify images of facial shapes, Human Face Shapes	High inaccurate classification results are produced	85,1 % of accuracy
18	Salim et al.[27]	2023	swin transformer model	To hardness the advantage of the swin transformer model for face shape classification	High inaccurate classification results are produced	By using data augmentation, 84,95 % precision, 87,37 % recall, 86,68 % F1-score, and 86,34 % accuracy
19	Srinivas et al.[38]	2023 (published in Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization Journal)	Inception V3 + GRU	To focus on face shape classification system using deep learning & multi-model feature fusion	The correlation-based hybrid feature fusion is not orthogonal which can cause poor discrimination ability	Accuracy of 98,51 %
20	Salim et al.[46]	2023	Swin transformer	To develop an advanced face shape classification system utilizing the Swin Transformer model.	Complex Model Architecture. Interpretability Issues & Risk of Overfitting	Accuracy of 86,34 %
21	Rabea et al.[47]	2024	VGG-16	To design and implement a robust multimodal facial biometric system called IdentiFace using the VGG neural network architecture.		Accuracy of 88,51 %
22	Srinivas et al.[37]	2025 (in final review with Expert Systems with Applications Journal)	Binary emperor penguin algorithm + honey badge bionic model	To develop face shape classification system based on deep learning with binary metaheuristic feature selection	Features are extracted from images only, hence status feature points alone employed for classification	Accuracy of 99 % with Face Shape Dataset
23	Srinivas et al.[36]	2025 (in final review with Computational Intelligence Journal)	GABF (Gaussian Bilateral Filtering) + HOG (Histogram of Oriented Gradients) + Swin-PCT Model + LSTM	To develop face shape classification system based on enhanced deep feature extraction and confined multi-model feature fusion	Not supported for 3D Face Shape Recognition	Accuracy of 98,96 % with Face Shape dataset and Labelled Face Dataset

Table 2. Women Face shape classification outcomes

Original image	Pre-processed image	Identification of facial landmarks and boundaries	Characteristics	Classification outcomes
			Balanced, Diplomat, Creative	Heart
			Deep thinkers, Value logic	Oblong
			Practical, Methodical	Oval
			Sensitive, Caring	Round
			Intelligent, Analytical, Decisive, Bold	Square

Source: outcome of MML-FuseNet: An Efficient Multi-Task Mutual Learning Framework for Face Shape Detection Based on Confined Feature Fusion. (2025) Computational Intelligence[36]



Source: Outcome of Deep Learning based Face Shape Classification System with Binary Feature Selection Model. (2025) *Expert Systems with Applications*[37]

Figure 19. Analysis of image outcome for men face shape dataset

DISCUSSION

This review highlights tremendous progress in the use of deep learning models for facial character assessment. High dimensionality of images is often addressed using convolutional neural networks (CNNs), which are exceedingly well at multiplying pixels (features) in the appropriate combinations based on local connections and shared weights, therefore allowing the algorithm to learn the regions around local pixels (features). CNNs are outperforming others in image classification and representing multiple images, while recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) are good at storing temporal dynamics of facial expressions. GANs plays an important role by creating synthetic facial data, which improves the dataset diversity and robustness. CNNs paired with RNNs or GANs in hybrid models have been shown to outperform previous image classification algorithms in tasks such as emotion detection or facial recognition. With the help of transfer learning, rapid model adaptation to new tasks with small amounts of data has been possible, leading to increased efficiency and accuracy for the models.

CONCLUSIONS

With the rise of deep learning, facial character study has made significant progress and improvement, including accuracy and robustness, as well as the generalization capability in various fields. Future directions include, but not limited to, developing systems for real-time processing, integrating multi-modal biometric, algorithms that can perform bio preservation with ethical nature. Tackling these challenges will be key to unlocking the full potential of facial recognition technology.

In comparison to existing models, the latest research paper model achieves a high accuracy of 98,96 % for the Face Shape Dataset and 98,88 % for the Labelled Face Dataset. These results underscore the model's effectiveness in accurately categorizing various facial shapes however this model has a limitation on 3D Face Shape features recognition.

Also, more efficient metaheuristic optimization algorithms can be developed in future works, leading to a reduction in feature dimension while improving accuracy. The adoption of more advanced metaheuristic optimization algorithms might facilitate a breakdown of feature dimensionality at the cost of degrading or improving accuracy. This would allow for more efficient models that are computationally cheaper. In addition, future work will focus on making the AI explainable, so that we understand their process of arriving to the predictions.

BILBIOGRAPHIC REFERENCES

1. Albiero V, Chen X, Yin X, Pang G, Hassner T. img2pose: Face alignment and detection via 6dof, face pose estimation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2021:7617-7627.

2. Wu Y, Ji Q. Facial landmark detection: A literature survey. International Journal of Computer Vision. 2019; 127:115-42.

3. Ji W, Jin L. Face Shape Classification Based on MTCNN and FaceNet. In 2021 2nd International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) 2021:167-170. IEEE.

4. Al-Bander B, Alzahrani T, Alzahrani S, Williams BM, Zheng Y. Improving fetal head contour detection by object localisation with deep learning. In Annual Conference on Medical Image Understanding and Analysis 2019:142-150. Cham: Springer International Publishing.

5. Saichandran KS. Hand Shape and Gesture Recognition using Multi-scale Template Matching, Background Subtraction and Binary Image Analysis. arXiv preprint arXiv:2402.09663. 2024.

6. Hossam M, Afify AA, Rady M, Nabil M, Moussa K, Yousri R, Darweesh MS. A comparative study of different face shape classification techniques. In 2021 International Conference on Electronic Engineering (ICEEM) 2021:1-6. IEEE.

7. Alzahrani T, Al-Nuaimy W, Al-Bander B. Hybrid feature learning and engineering based approach for face shape classification. In 2019 International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS) 2019:1-4. IEEE.

8. Tio AE. Face shape classification using inception v3. arXiv preprint arXiv:1911.07916. 2019.

9. Marinescu AI, Ileni TA, Darabant AS. A versatile 3d face reconstruction from multiple images for face shape classification. In 2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM) 2019:1-6. IEEE.

https://doi.org/10.56294/dm2025658

10. Razalli H, Rahmat RW, Khalid F, Sulaiman PS. An Image-Based Children Age Range Verification and Classification Based on Facial Features Angle Distribution and Face Shape Elliptical Ratio. Advanced Science Letters. 2017; 23(5):4026-30.

11. Deng J, Trigeorgis G, Zhou Y, Zafeiriou S. Joint multi-view face alignment in the wild. IEEE Transactions on Image Processing. 2019; 28(7):3636-48.

12. Nabil M, Rady M, Moussa K, Wessam M, Hossam M, Yousri R, Darweesh MS. A pre-processing approach to improve the performance of inception v3-based face shape classification. In 2021 3rd Novel Intelligent and Leading Emerging Sciences Conference (NILES) 2021:205-209. IEEE.

13. Sabir E, Cheng J, Jaiswal A, AbdAlmageed W, Masi I, Natarajan P. Recurrent convolutional strategies for face manipulation detection in videos. Interfaces (GUI). 2019; 3(1):80-7.

14. Abdullah, Hussain A, Ali S, Kim HC, Sain M, Aich S. Hybrid Based Model Face Shape Classification Using Ensemble Method for Hairstyle Recommender System. In International conference on smart computing and cyber security: strategic foresight, security challenges and innovation 2021:61-68. Singapore: Springer Nature Singapore.

15. Polat S, Kabakc A, Cevik Y, YÜCEL A. The Face Shape and Golden Ratio Classification in Turkish Healthy Population. Journal of Evolution of Medical and Dental Sciences-Jemds. 2020; 9(2).

16. Zhou D, Jin X, Jiang Q, Cai L, Lee SJ, Yao S. MCRD-Net: An unsupervised dense network with multi-scale convolutional block attention for multi-focus image fusion. IET Image Processing. 2022; 16(6):1558-74.

17. Thevenot J, López MB, Hadid A. A survey on computer vision for assistive medical diagnosis from faces. IEEE journal of biomedical and health informatics. 2017; 22(5):1497-511.

18. Rahmat RF, Syahputra MD, Andayani U, Lini TZ. Probabilistic neural network and invariant moments for men face shape classification. In IOP Conference Series: Materials Science and Engineering 2018; 420(1):012095. IOP Publishing.

19. Ross A, Othman A. Visual cryptography for biometric privacy. IEEE transactions on information forensics and security. 2010; 6(1):70-81.

20. Rossler A, Cozzolino D, Verdoliva L, Riess C, Thies J, Nießner M. Faceforensics++: Learning to detect manipulated facial images. In Proceedings of the IEEE/CVF international conference on computer vision 2019:1-11.

21. Mehta A, Mahmoud T. Human Face Shape Classification with Machine Learning.

22. Sukumaran A, Brindha T. Optimal feature selection with hybrid classification for automatic face shape classification using fitness sorted Grey wolf update. Multimedia Tools and Applications. 2021; 80:25689-710.

23. Alzahrani T, Al-Nuaimy W, Al-Bander B. Integrated multi-model face shape and eye attributes identification for hair style and eyelashes recommendation. Computation. 2021; 9(5):54.

24. Zhao J, Zhang M, He C, Xie X, Li J. A novel facial attractiveness evaluation system based on face shape, facial structure features and skin. Cognitive Neurodynamics. 2020; 14:643-56.

25. Shao Z, Liu Z, Cai J, Ma L. Jaa-net: joint facial action unit detection and face alignment via adaptive attention. International Journal of Computer Vision. 2021; 129:321-40.

26. Adityatama R, Putra AT. Image classification of Human Face Shapes Using Convolutional Neural Network Xception Architecture with Transfer Learning. Recursive Journal of Informatics. 2023; 1(2):102-9.

27. Salim BV, Indrawan JO, Hidayat J, Matthew S, Mangkang TA, Hasana S, Permonangan IH. Face Shape Classification Using Swin Transformer Model. Procedia Computer Science. 2023; 227:557-62.

28. VenkateswarLal P, Nitta GR, Prasad A. Ensemble of texture and shape descriptors using support vector machine classification for face recognition. Journal of Ambient Intelligence and Humanized Computing. 2019:1-8.

29. Radhika R, Mahajan R. An adaptive optimum weighted mean filter and bilateral filter for noise removal in cardiac MRI images. Measurement: Sensors. 2023; 29:100880.

30. Rajani Kumari LV, Saher Fathima S, Sai Praneeth G, Mamatha D, Pranitha B. Dynamic face recognition system using histogram of oriented gradients and deep neural network. InSustainable Communication Networks and Application: Proceedings of ICSCN 2021 2022 (pp. 229-241). Singapore: Springer Nature Singapore.

31. Goodfellow, I., et al. (2014). Generative Adversarial Networks. Advances in Neural Information Processing Systems.

32. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.

33. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems.

34. LeCun, Y., et al. (1998). Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE, 86(11), 2278-2324.

35. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. IEEE Conference on Computer Vision and Pattern Recognition.

36. Srinivas, A. & Vamsidhar, E., (2025). MML-FuseNet: An Efficient Multi-Task Mutual Learning Framework for Face Shape Detection Based on Confined Feature Fusion. Computational Intelligence.

37. Srinivas, A. & Vamsidhar, E., (2025). Deep Learning based Face Shape Classification System with Binary Feature Selection Model. Expert Systems with Applications.

38. Srinivas, A. & Vamsidhar, E., (2023). Multimodal face shape detection based on Human Temperament with Hybrid Feature Fusion and Inception V3 Extraction Model. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization.

39. Zafar, A., & Popa, T. (2016). Face and eye-ware classification using geometric features for a data-driven eye-ware recommendation system. In Graphics Interface Conference (pp. 183-188). New York, NY, USA: ACM Library.

40. Sarakon, P., Charoenpong, T., & Charoensiriwath, S. (2014, November). Face shape classification from 3D human data by using SVM. In The 7th 2014 Biomedical Engineering International Conference (pp. 1-5). IEEE.

41. Bansode, N. K., & Sinha, P. K. (2016). Face shape classification based on region similarity, correlation and fractal dimensions. International Journal of Computer Science Issues (IJCSI), 13(1), 24.

42. Sunhem, W., & Pasupa, K. (2016, February). An approach to face shape classification for hairstyle recommendation. In 2016 Eighth International Conference on Advanced Computational Intelligence (ICACI) (pp. 390-394). IEEE.

43. Rahmat, R. F., Syahputra, M. D., Andayani, U., & Lini, T. Z. (2018, September). Probabilistic neural network and invariant moments for men face shape classification. In IOP Conference Series: Materials Science and Engineering (Vol. 420, No. 1, p. 012095). IOP Publishing.

44. Tio, A. E. (2019). Face shape classification using inception v3. arXiv preprint arXiv:1911.07916.

45. Pasupa, K., Sunhem, W., & Loo, C. K. (2019). A hybrid approach to building face shape classifier for hairstyle recommender system. Expert Systems with Applications, 120, 14-32.

46. Salim, B. V., Indrawan, J. O., Hidayat, J., Matthew, S., Mangkang, T. A. E., Hasana, S., & Permonangan, I. H. (2023). Face Shape Classification Using Swin Transformer Model. Procedia Computer Science, 227, 557-562.

47. Rabea, M., Ahmed, H., Mahmoud, S., & Sayed, N. (2024). IdentiFace: A VGG Based Multimodal Facial Biometric System. arXiv preprint arXiv:2401.01227.

https://doi.org/10.56294/dm2025658

FINANCING
No financing: the authors did not receive financing for the development of this research.

CONFLICT OF INTEREST
None: the authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION
Data curation: Srinivas Adapa, Vamsidhar Enireddy.
Methodology: Srinivas Adapa, Vamsidhar Enireddy.
Software: Srinivas Adapa, Vamsidhar Enireddy.
Drafting - original draft: Srinivas Adapa, Vamsidhar Enireddy.
Writing - proofreading and editing: Srinivas Adapa, Vamsidhar Enireddy.