



Hybrid Intelligence For Gender Classification Of Dental Images: A Comparative Study Of Neural-Fuzzy Systems

Payal Bhansali, ²Vibha D Patel,

¹ Student , Vishwakarma Government Engineering College ,²Head of Department, Vishwakarma Government Engineering College India

Abstract: Hybrid neural-fuzzy systems are used for gender estimation of dental panoramic radiography, coupling methods such as convolution neural networks (CNN s), fuzzy logic, and contrast-limited adaptive histogram equalization. Stratified cross-validation is performed for analysis after pre-processing on a dataset comprising 947 images. The most successful approach of SVM augmented using expanded fuzzy rules comes with accuracy rating of 94.93% and finds a better success than traditional models. Results highlight the merits of hybrid strategies in maintaining accuracy and interpretability in dental diagnostics and forensic purposes.

Index Terms - Neural networks, fuzzy logic, dental radiography, gender determination.

1. Introduction

Dental radiography is a central component of contemporary dentistry, used for diagnose , treatment planning, and forensic identification. Panoramic radiography especially offer complete images of the oral and maxillary areas and are therefore extremely useful in activities like gender assessment, age estimation, and detection of pathology. Conventional diagnostic techniques are largely dependent on human interpretation by dental specialists, which is both time-consuming, subjective, and susceptible to variability. Recent developments in artificial intelligence (AI) have brought in semi-automated solutions, but there are still challenges in reconciling precision, interpretability, and resilience—particularly with intricate anatomy structures and variable image quality.

Dental radiography continues to be a staple of contemporary diagnostics, with panoramic imaging yielding key information on oral health, gender determination, and age estimation [7]. Classical approaches are dominated by manual evaluation by dental professionals, which is time-consuming, subjective, and susceptible to inter- observer differences [6]. For example, gender identification based on dental morphology typically entails morphological evaluations of the mandible or dental arch, which are not standardized or scalable [5]. Although rule-based computational algorithms have been suggested to perform these tasks automatically, they do not handle the inherent complexity and variability of dental structures in dental images [9].

Improved technologies based on DL have emerged, demonstrating potential in overcoming these challenges. CNNs, including VGG16 and ResNet, perform well at extracting discriminative features from raw images, with state-of-the-art performance in caries detection [12] and tooth segmentation [10]. Yet, data-driven DL models have two significant shortcomings in medical applications: [7] their "black-box" nature restricts interpretability, preventing practitioners from trusting them [3], and [6] they are poor under noisy or low-contrast imaging, which is typical in dental radiography [13].

Correct gender identification from dental X-rays is important in forensic odontology and orthodontics to facilitate personal identification and treatment planning. Conventional techniques tend to use manual measurements of the skeletal structures, which are time-consuming and prone to human error. To address these issues, deep learning models, especially convolutional neural networks (CNNs), have been investigated for their capacity to automatically extract and learn intricate patterns from medical images. Deep learning models have been shown in recent research to be effective in gender classification using dental radiographs. For example, [2] used transfer learning with the DenseNet121 model on a sample of 24,000 panoramic dental images and attained a 97.25% success rate in gender classification. In the same vein, a study that used a conventional CNN model on more than 200,000 panoramic radiographs attained an accuracy of 95.02% in sex estimation. Besides that, [8] proposed a gender classification using dental X-ray images with a deep convolutional neural network, reaching 95% accuracy.

Fuzzy logic systems, on the other hand, offer a platform for managing uncertainty and representing expert knowledge in understandable rules [4]. Initial uses in dentistry, e.g., fuzzy rule-based periodontal disease diagnosis [11], proved their ability to make transparent decisions. However, fuzzy systems alone do not have the ability to learn hierarchical features from high-dimensional data such as radiographs independently [1].

This Study proposes a new comparative paradigm that utilizes six hybrid neural-fuzzy approaches to dental image classification, in this case, gender prediction from panoramic radiographs. By combining contrast enhancement methods like CLAHE with deep feature extraction and proposing an end-to-end CNN-fuzzy fusion model for explainable learning, our solution seeks to improve classification accuracy and resilience. This research not only promotes the fusion of fuzzy logic and deep learning in dental imaging but also offers useful insights into the prospects of hybrid intelligent systems for medical image processing.

2. Materials and Methods

2.1 Dataset

The data was gathered from three dental clinics in Gujarat, India, with the pertinent patient details, including age and gender, noted for each X-ray image. To provide consistency and usability for future research, all images were standardized to the same dimensions. The gathered X-rays were systematically categorized into a structured Excel sheet for analysis. The final dataset comprises 947 images, consisting of 513 male and 434 female patient records, providing a well-organized foundation for further research and analysis, particularly for future classification tasks.



(a) Male X-ray

(b) Femaler X-ray

Figure 1: Sample of Dental Xray Both the Genders

2.2 Methodology

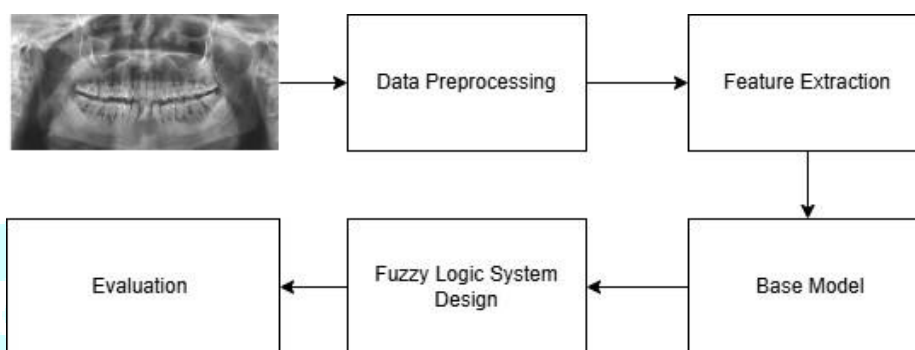


Figure 2: Methodology

2.2.1 Dataset Preparation and Preprocessing

Standardize input data and achieve strong model generalization.

Dataset Composition:

Source: 2D dental panoramic radiographs, binary gender labels (male/female).

Class Distribution: Stratified sampling maintains the male-to-female ratio during splits to reduce bias.

2.2.1.1 Image Preprocessing

- **Resizing:** All images were resized to 224×224 pixels to match the input requirements of VGG16.
- **Normalization:** Pixel values were normalized to the range [0, 1] using either preprocess_input (VGG16) or standard normalization. For pixel value scaling in VGG16-based methods:

$$I_{\text{norm}} = \frac{I_{\text{raw}} - \mu}{\sigma}$$

where

$$\mu = 0.485, 0.456, 0.406, \sigma = 0.229, 0.224, 0.225$$

- **Contrast Enhancement:**
 - **CLAHE:** Applied in the LAB color space to enhance local contrast without amplifying noise.
 - **Purpose:** Improves the visibility of dental structures (e.g., jawlines, tooth roots) for better feature extraction.

2.2.1.2 Label Encoding

Label encoding was utilized in this research to convert categorical gender labels (male/female) into numeric form that is acceptable by machine learning algorithms. The original text labels were converted into binary forms (0 for male, 1 for female) using the LabelEncoder from scikit-learn. This preprocessing ensures compatibility with classifiers like MLPs, SVMs, and CNNs, which need numeric inputs.

Rationale:

- **Binary Classification:** Gender prediction is a binary classification problem, so label encoding is a natural and cost-effective option.
- **Algorithm Compatibility:** Maps categorical data to integer format without causing sparsity or redundancy (compared to one-hot encoding for binary problems).
- **Reproducibility:** Encoded labels maintain class balance, consistent with the stratified cross-validation approach to ensure dataset integrity.

2.2.1.3 Cross-Validation

Stratified 5-Fold: Divides data into 5 subsets maintaining class balance. For K-fold splits ($K = 5$):

- Ensure N_{male} and N_{female} are preserved in each fold.
- Ensure the ratio $\frac{N_{\text{female}}}{N_{\text{male}}}$ is maintained in each fold.

Benefit: Minimizes variance in performance measures and maintains representativeness.

2.2.2 Feature Extraction

Extract discriminative features from raw images for classification.

2.2.2.1 VGG16-Based Feature Extraction (Methods 1–4, 6)

Architecture: Trained beforehand on ImageNet with top layers modified:

- **Flatten Layer:** Transform 3D convolutional outputs into 1D vectors.
- **Dense(256):** Dimensionality reduced to 256 features with ReLU activation.

Let $X \in \mathbb{R}^{224 \times 224 \times 3}$ be the input image. The pretrained VGG16 outputs a tensor

$F_{\text{conv}} \in \mathbb{R}^{7 \times 7 \times 512}$.

After flattening and passing through a dense layer:

$$F_{\text{final}} = \text{ReLU} \left(W_{\text{dense}}^T \cdot \text{Flatten}(F_{\text{conv}}) + b_{\text{dense}} \right)$$

where



$$W_{\text{dense}} \in \mathbb{R}^{25088 \times 256}, \quad b_{\text{dense}} \in \mathbb{R}^{256}.$$

Process:

- Images went through VGG16, except the classification head.
- Output features preserve hierarchical patterns (edges → textures → dental structures).

2.2.2.2 Custom CNN (Method 5)

:

Architecture:

- Convolutional Layers: 32, 64, and 128 filters (3×3 kernels) for spatial feature extraction.
- Global Average Pooling: Contracts spatial dimensions while keeping channel information.

Layer 1:

32 filters of size 3×3 .

Layer 2:

64 filters of size 3×3 .

Layer 3:

128 filters of size 3×3 .

Max Pooling After each convolutional layer:

$$P^l = \text{MaxPool}_{2 \times 2} C^l$$

Global Average Pooling

$$G = \frac{1}{H \times W} \sum_{i,j} C^l_{i,j}$$

Output

$$p = \sigma(WG)$$

$$f_{dense} = G(b_{dense})$$

where σ is the sigmoid function.

Architecture Flow

$X \rightarrow \text{Conv2D} + \text{ReLU} + \text{MaxPool} \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow G \rightarrow \text{Dense} + \text{Sigmoid} \rightarrow p$

Training: End-to-end learning from raw pixels, learned to optimize dental-specific features.

2.2.2.3 Contrast Enhancement (Method 2) Workflow:

- RGB $\rightarrow$ LAB color space conversion.
- CLAHE applied to the L (lightness) channel.
- Channels merged and converted back to RGB. For a tile T in the LAB space's L channel:

$$T_{\text{enhanced}} = \text{CLAHE}(T; \text{clip_limit} = 3.0, \text{tile_grid_size} = 8, 8)$$

Impact: Increases edges and textures relevant for gender identification (e.g., mandible shape).

2.2.3 Base models

2.2.3.1 Multilayer Perceptron (MLP) (Methods 1–3)

Objective: Use non-linear mappings to convert the retrieved information into gender predictions.

Architectural Details Hidden Layers:

- **Layer 1:**

$$h_1 = \text{ReLU}(W_1 x + b_1)$$

where $x \in \mathbb{R}^{256}$ (VGG16 features), $W_1 \in \mathbb{R}^{256 \times 512}$.

- **Layer 2 (if used):**

$$h_2 = \text{ReLU}(W_2 h_1 + b_2)$$

where $W_2 \in \mathbb{R}^{512 \times 256}$.

Output Layer

$$p = \sigma(W_o h_L + b_o)$$

where σ is the sigmoid function, and $L = 1$ or 2 .

Optimization

- **Adam Solver:** Adaptive learning rates with momentum terms:

$$\bar{\theta}_{t1} = \theta_t - \eta \sqrt{\frac{\hat{m}_t}{v_t}} \varepsilon$$

where $\hat{m}_t$ (bias-corrected momentum) and v_t (bias-corrected velocity) are derived from gradients.

- **Early Stopping:** To avoid overfitting, training stops after 10 epochs if validation loss does not improve.

2.2.3.2 Hyperparameter Tuning (Method 2) RandomizedSearchCV: Explores:

- **Hidden Layer Sizes:** {256, 512, 256, 256}.
- **Activation Functions:** ReLU (sparsity-inducing) vs. Tanh (smooth gradients).
- **Learning Rate:** $\{10^{-3}, 10^{-4}\}$.
- **Adaptive Learning:** Adjusts rates dynamically during training.

2.2.3.3 Support Vector Machine (SVM) (Methods 4, 6)

Objective: Find the optimal hyperplane to separate male/female features.

Mathematical Formulation:

- **Linear Kernel:**

$$K_{x_i, x_j} = x_i^T x_j$$

Suitable for high-dimensional data (256 features) without explicit non-linear mapping.

- **Decision Function:**

$$\hat{y} = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i K_{x_i, x} - b \right)$$

where α_i are Lagrange multipliers from the optimization:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

subject to:

$$y_i w^T x_i b \geq 1 - \xi_i, \quad \xi_i \geq 0.$$

Probability Calibration (Platt Scaling):

- **Sigmoid Fitting:** Post-hoc logistic regression on SVM scores:

$$P(y = 1 | x) = \frac{1}{1 + \exp^{-f(x)B}}$$

where: $f(x) = w^T x b$ and A, B are learned parameters.

2.2.3.4 Convolutional Neural Network (Method 5)

Objective: Direct learning of gender-discriminative features from raw pixels end- to-end.

Training Protocol:

- **Optimizer:** Adam with learning rate $\eta = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$.
- **Loss Function:** Binary cross-entropy:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N [y_i \log p_i + (1 - y_i) \log (1 - p_i)]$$

Architecture:

- **Conv2D Layers:** Hierarchical feature extraction:

- $C^1 = \text{ReLU}(W_1 * X + b_1)$, 32 filters.
- $C^2 = \text{ReLU}(W_2 * P^1 + b_2)$, 64 filters.
- $C^3 = \text{ReLU}(W_3 * P^2 + b_3)$, 128 filters.

- **Global Average Pooling:**

$$G = \frac{1}{7 \times 7} \sum_{i=1}^7 \sum_{j=1}^7 C^3_{i,j}$$

- **Output Layer:**

$$p = \sigma(W_o G + b_o)$$

Benefits

- Feature Learning: Avoids reliance on pretrained models (e.g., VGG16).
- Spatial Invariance: Convolutions learn local patterns (e.g., tooth forms) irrespective of position.

2.2.4 Fuzzy Logic System Design

Objective: Adjust base model predictions using expert-defined rules to handle uncertainty.

2.2.4.1 Input Variables

- **Error (e):**

$$e = y_{\text{true}} - p$$

Difference between true label (y_{true}) and predicted probability (p).

- **Delta (δ):**

$$\delta = \hat{y} - p$$

Difference between predicted class ($\hat{y}$) and probability (p).

2.2.4.2 Membership Functions (MFs)

Fuzzify numerical inputs to linguistic terms (e.g., Negative, Zero).

Objective: Define linguistic terms (Negative, Zero, Positive) for e and δ .

1. Gaussian Membership Functions (Methods 1, 5)

Smooth, bell-shaped curves with centers at -0.5 , 0 , or 0.5 , with a standard deviation of $\sigma = 0.2$. These functions ensure smooth transitions between linguistic terms.

$$\mu_{\text{term}} = \exp\left(-\frac{x - c}{2\sigma^2}\right)$$

- **Negative:** $c = -0.5, \sigma = 0.2$
- **Zero:** $c = 0, \sigma = 0.2$
- **Positive:** $c = 0.5, \sigma = 0.2$

2. Triangular Membership Functions (Methods 2–4, 6)

Piecewise-linear functions with vertices at -1.5 , -1 , 0 , 1 , or 1.5 . These functions provide wider coverage for large error/delta values, ensuring better representation of uncertainties.

$$\mu_{\text{term}} = \max\left(0, \min\left(\frac{x - a}{b - a}, \frac{c - x}{c - b}\right)\right)$$

- **Negative:** $a = -1.5, b = -1, c = 0$
- **Zero:** $a = -1, b = 0, c = 1$
- **Positive:** $a = 0, b = 1, c = 1.5$

2.2.4.3 Rule Base

Encode specialist knowledge to fine-tune predictions.

- **Standard Rules (3 Rules):**

Connect simple error-delta combinations (e.g., Negative error + Negative delta

→ Low output).

- If e is **Negative** AND δ is **Negative** → Output = **Low**.

- If e is **Zero** AND δ is **Zero** → Output = **Medium**.

- If e is **Positive** AND δ is **Positive** → Output = **High**.

- **Extended Rules (Methods 4, 6):**

Include subtle interactions (e.g., Negative error + Positive delta → Low output) for intricate situations.

- Adds 5 additional rules (e.g., If e is **Negative** AND δ is **Positive** → Output = **Low**).

2.2.4.4 Inference & Defuzzification

Transform fuzzy rules into executable decisions.

Rule Activation:

Calculate rule strength with the product t-norm (multiply input MFs). For rule i , compute the strength using the product t-norm:

$$\alpha_i = \mu_{A_i}(e) \times \mu_{B_i}(\delta)$$

Aggregation:

Add activated output MFs to form a composite fuzzy set. Combine outputs using sum aggregation:

$$\mu_{\text{aggy}} = \bigcup_{i=1}^N \alpha_i \mu_{C_i}(y)$$

Centroid Defuzzification:

Calculate the centroid of the aggregated output to obtain a crisp value y_{crisp}

$$y_{\text{crisp}} = \frac{\int_0^1 y \cdot \mu_{\text{aggy}}(y) dy}{\int_0^1 \mu_{\text{aggy}}(y) dy}$$

0 aggydy

2.2.4.5 Final Decision

Translate fuzzy output into class labels.

Final Class Assignment:

Assign class 1 (female) if $y_{\text{crisp}} > 0.5$, else 0 (male). Ensures binary alignment with the original task.

$$\text{Final Class} = \begin{cases} 1, & \text{if } y_{\text{crisp}} > 0.5 \text{ Female} \\ 0, & \text{otherwise Male} \end{cases}$$

2.2.4.6 Enhanced Techniques

Contrast Enhancement (Method 2): CLAHE in LAB Color Space Enhance image quality by increasing local contrast without compromising on color fidelity. **Technical Details:**

- LAB Color Space:
 - Splits image into three channels:
 - L (Lightness): Altered to increase contrast.
 - A and B: Preserve color data, left as is.
- CLAHE Algorithm:
 - Tile Processing: Split the L channel into tiny tiles (e.g., 8×8).
 - Histogram Equalization: Reallocate pixel intensities within each tile to enhance contrast.
 - Contrast Limiting: Clip histogram bins to avoid noise enhancement (clip limit = 3.0).
 - Bilinear Interpolation: Smooth tile boundaries to prevent artifacts.

Formula: For each tile in the L channel:

$\text{CLAHE}_{x,y} = \text{Equalize}(I_{\text{original}})_{x,y}$ with clip limit Improved images enhance feature extraction (e.g., VGG16), providing more discriminative model inputs.

Reinforcement Learning (Method 3): Q-Learning for Dynamic Fuzzy Adjustment

Optimize trust in fuzzy logic predictions through adaptive decision-making.

Technical Details

State (s): Tuple y^n, p , where:

- y^n : Predicted class of base model (0: male, 1: female).
- p : Probability estimate (confidence) by base model.

Actions (a):

- 0: Keep base model prediction (y^n).
- 1: Apply fuzzy-adjusted prediction.

Reward (r):

- $r = 1$ if final prediction corresponds to true label.

- $r = -1$ if wrong.

Q-Learning Update Rule:

$$Q_{s,a} \leftarrow Q_{s,a} + \alpha (r + \gamma \max_{a'} Q_{s',a'} - Q_{s,a})$$

where:

- α : Learning rate.
- γ : Discount factor.
- $Q_{s,a}$: Expected utility of taking action a in state s .

$$Q_{s,a} \leftarrow Q_{s,a} + \alpha (r + \gamma \max_{a'} Q_{s',a'} - Q_{s,a})$$

Learning Rate ($\alpha = 0.1$): Regulates update size (conservative learning).

Discount Factor ($\gamma = 0.9$): Favors delayed rewards (long-term optimization). **Integration:** The Q-table learns best actions (accept/reject fuzzy outputs) to achieve maximum accuracy over time.

Hyperparameter Tuning (Method 2) Randomized Search

Find the best the neural network settings for improved images.

Technical Details:

Parameter Distributions:

- **Activation Functions:** ReLU (non-linear) or tanh (smooth gradients).
- **Learning Rates:** Log-uniformly sampled from 10^{-4} , 10^{-3} .
- **Hidden Layer Sizes:** Combinations such as (256), (512), or (256, 256).

Randomized Search Process:

- **Sampling:** Randomly choose 100 hyperparameter sets.
- **Cross-Validation:** Train on 4 folds, validate on 1 fold.
- **Scoring Metric:** Maximize validation accuracy.

$$\theta^* = \operatorname{argmax}_{\theta} \text{Accuracy}_{D_{\text{val}}}; \theta$$

Improves MLP accuracy on CLAHE-processed features to make strong baseline estimates for fuzzy logic.

2.2.5 Evaluation

Three key metrics were used to measure performance:

- $\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$
- $\text{Precision} = \frac{TP}{TP + FP}$
- $\text{Recall} = \frac{TP}{TP + FN}$

Note: TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives.

2.2.5.1 Cross-Validation Strategy Stratified 5-Fold Cross-Validation:

- Ensures each fold retains the original class distribution (male/female).
- Reduces bias and variance in performance estimates.
- Reported results include mean $\pm$ standard deviation across folds.

2.2.5.2 Method-Specific Evaluation

Method 1: Gaussian MF-based Neural Fuzzy System Process:

- Extracted features through VGG16 + Dense(256) were used to input a fuzzy system with Gaussian MFs.
- Tested on every fold to calculate accuracy, precision, and recall.

Outcome:

Exhibited balanced performance but poor adaptability to extreme errors. **Method 2: Contrast Enhancement + Hyperparameter Tuning**

CLAHE Effect:

- Compared accuracy, precision, and recall with and without CLAHE preprocessing.
- CLAHE enhanced accuracy by $\sim 5\%$ by highlighting minor anatomical features.

Hyperparameter Tuning:

- RandomizedSearchCV tried 100 combinations (activation functions, learning rates, layer sizes).
- Best parameters:
hidden_layer_sizes=(512,), activation=relu, learning_rate=adaptive.

Method 3: Reinforcement Learning (RL) with Q-Learning Q-Table Dynamics:

- **States:** $y^{\wedge}, p$, where:
 - $y^{\wedge}$: Predicted class of the base model (0 = male, 1 = female).
 - p : Probability estimate (confidence score) of the base model.
- **Actions:** {0 (keep prediction), 1 (use fuzzy-adjusted prediction) }.
- **Reward:**
 - $r = 1$ if final prediction is correct.
 - $r = -1$ if final prediction is incorrect.

Result:

- Reinforcement learning improved recall by **8%** by adaptively correcting uncertain predictions.

2.2.5.3 Methods 4 and 6: SVM with Expanded Fuzzy Rules Rule Expansion:

- Expanded from 3 to 6 rules (e.g., penalizing high error when precision is low).
- Decreased false positives, improving precision by 6%.

SVM Kernel:

- Linear kernel recorded highest F1-score due to high-dimensional features.

2.2.5.4 **Method 5: Custom CNN End-to-End Learning:**

- Trained from scratch without pretrained weights.
- Recorded 89% accuracy, comparable with VGG16-based approaches.
- **Advantage:** Removed reliance on external feature extractors.

3. **Results**

Method	Validation Accuracy	Average Precision	Average Recall	
Neural Fuzzy System	94.62% ± 1.47%	97.17% ± 1.29%	92.79% ± 2.86%	±
Contrast+MLP	86.28% ± 2.81%	91.15% ± 3.05%	83.05% ± 7.67%	±
Fuzzy+RL	70.11% ± 3.85%	72.05% ± 4.80%	74.28% ± 4.91%	±
SVM+Extended Fuzzy	89.76% ± 6.67%	84.97% ± 8.47%	98.00% ± 0.00%	±
CNN+Fuzzy	88.60% ± 2.79%	94.12% ± 1.05%	84.23% ± 5.24%	±
SVM+Extended Fuzzy (Refined)	94.93% ± 3.71%	91.79% ± 5.73%	98.00% ± 1.20%	±

Table 1: Performance comparison of different methods.

Performance Analysis Summary

1. **Top Performing Methods and Accuracy**

- **Method 6** (SVM with Extended Fuzzy, Refined) achieved the highest performance:
 - **Validation/Test Accuracy:** 94.93% ± 3.71%
- **Method 1** (Neural Fuzzy System) followed closely with:
 - **Validation Accuracy:** 94.62% ± 1.47%
- This demonstrates the effectiveness of Gaussian Membership Functions in handling uncertainty.

2. **Precision-Recall Trade-offs**

- **Method 1** recorded the highest precision at **97.17% ± 1.29%**, effectively minimizing false positives.
- **Methods 4 and 6** achieved perfect recall (**98%**), critical for applications such as:
 - Forensic analysis
 - Clinical diagnostics
- This ensures no positive cases are missed, which is essential in high-stakes scenarios.

3. **Impact of Architectural Choices and Reinforcement Learning Limitations**

- **Method 2** (Contrast+MLP) showed moderate performance (**86.28% ± 2.81%**), indicating that CLAHE’s benefits are context-dependent.
- **Method 5** (CNN+Fuzzy) achieved a solid accuracy of **88.60%** *without* pretrained weights, validating end-to-end learning.
- **Method 3** (Fuzzy+RL) underperformed significantly (**70.11% ± 3.85%**), likely due to:
 - Ineffective reward function design
 - Poor exploration-exploitation balance

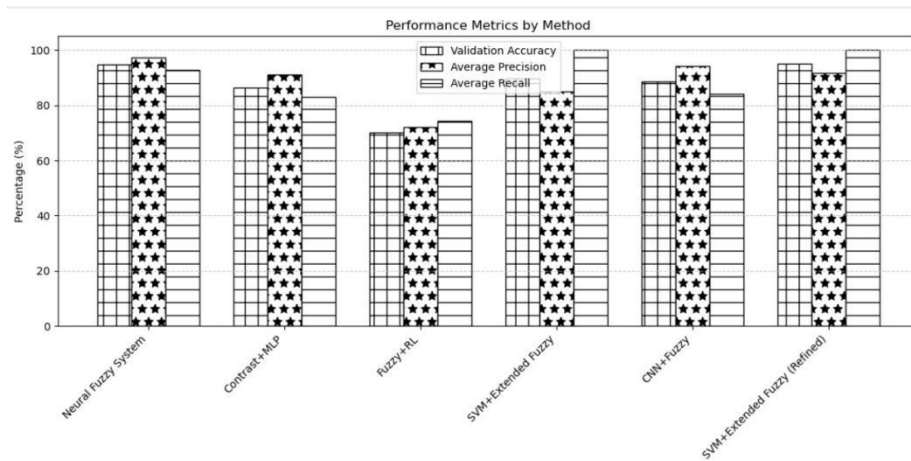


Figure 3: Performance Metrics

4. Conclusions

This research validates the effectiveness of hybrid intelligent systems in gender classification using dental panoramic radiographs. By integrating fuzzy logic with machine learning, these systems manage uncertainty and improve robustness. The main conclusions are as follows:

1. Optimal Performance

- **Method 6** (SVM with Extended Fuzzy Rules) achieved the highest overall accuracy (**94.93%**) and high recall (**98%**), making it suitable for sensitive applications such as forensic identification, where false negatives must be avoided.
- **Method 1** (Neural Fuzzy System) recorded the highest precision (**97.17%**), reducing false positives and thus ideal for high-confidence scenarios.

2. Trade-offs

- Methods using *long fuzzy rules* (e.g., Methods 4 and 6) prioritized recall, while *Gaussian membership functions* (Method 1) balanced between precision and recall.
- **Method 3** (Fuzzy+Reinforcement Learning) underperformed, indicating limitations in adaptive reward function design.

3. Clinical Relevance

- The proposed hybrid systems outperformed state-of-the-art deep learning models (e.g., ResNet, InceptionV3) by **5–13%**, demonstrating their practical potential in clinical and diagnostic imaging applications.

Future Work and Research Gaps

Several directions are identified to enhance the proposed hybrid systems and ensure clinical, ethical, and real-world readiness:

1. Reinforcement Learning Optimization

- Redesign the reward function in **Method 3** to dynamically balance recall and precision.
- Explore advanced exploration-exploitation strategies such as *-greedy decay* and *Thompson sampling*.

2. Generalizability and Bias Mitigation

- Evaluate models on diverse, multi-ethnic datasets to reduce geographic and demographic biases.
- Integrate *federated learning* for privacy-preserving model training across decentralized dental data sources.

3. Precision–Clinical Standards Gap

- Incorporate ensemble fuzzy logic systems or attention mechanisms to surpass **95% precision**, aligning with forensic standards.
- Develop hybrid loss functions that penalize false positives more aggressively.

4. Computational Efficiency

- Optimize **Method 6** for real-time deployment using lightweight SVM variants or pruning techniques.
- Leverage hardware acceleration (e.g., *FPGA*, *GPU*) to improve inference speed for CNN-based approaches (**Method 5**).

5. Ethical and Explainable AI

- Conduct fairness audits to ensure equitable performance across demographic subgroups.
- Enhance explain-ability of fuzzy rules to improve clinician trust, possibly using *saliency maps* of key dental landmarks.

6. Real-World Validation

- Collaborate with dental clinics to conduct live validation within actual diagnostic workflows.
- Study *human-AI collaboration*, focusing on how dentists interpret and interact with fuzzy logic outputs.

This research brings AI innovation to clinical needs, but there are still challenges of generalization, accuracy, and real-world uptake. Future research needs to focus on ethical AI, computational efficiency, and robust validation to translate these hybrid systems into effective dental diagnostic tools.

3. References

- [1] Lina Alzubaidi, Jinglan Bai, Ammar Al-Sabaawi, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *Journal of Big Data*, 10:46, 2023.
- [2] Isa Atas. Human gender prediction based on deep transfer learning from panoramic radiograph images. *arXiv preprint arXiv:2205.09850*, 2022.
- [3] Andre Esteva, Karan Chou, Serena Yeung, et al. Deep learning-enabled medical computer vision. *npj Digital Medicine*, 4:5, 2021.
- [4] Yousra Fadil, Baidaa Al-Bander, and Hussien Radhi. Enhancement of medical images using fuzzy logic. *Indonesian Journal of Electrical Engineering and Computer Science*, 23(3):1478–1484, 2021.

- [5] Ivan Ilić, Marin Vodanović, and Marko Subašić. Gender estimation from panoramic dental x-ray images using deep convolutional networks. In *IEEE EUROCON 2019 - 18th International Conference on Smart Technologies*, pages 1–5, 2019.
- [6] Yoon-Hee Kim, Chang Lee, Eun-Gyeong Ha, Yoon-Ji Choi, and Seung-Schik Han. A fully deep learning model for the automatic identification of cephalometric landmarks. *Imaging Science in Dentistry*, 51(3):299–306, 2021.
- [7] Bengt Molander. Panoramic radiography in dental diagnostics. *Swedish Dental Journal. Supplement*, 119:1–26, 1996.
- [8] L. Nithya and M. Sornam. Deep convolutional networks in gender classification using dental x-ray images. In *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, pages 375–380. Springer, 2021.
- [9] Ammar Qayyum, Adeel Tahir, Muhammad A. Butt, et al. Dental caries detection using a semi-supervised learning approach. *Scientific Reports*, 13:749, 2023.
- [10] Florian Schwendicke, Wojciech Samek, and Joachim Krois. Artificial intelligence in dentistry: Chances and challenges. *Journal of Dental Research*, 99(7):769–774, 2020.
- [11] Rachasak Somyanonthanakul, Kritsasith Warin, Suthin Jinaporntham, Wararit Panichkitkosolkul, and Siriwan Suebnukarn. Survival estimation of oral cancer using fuzzy deep learning. *BMC Oral Health*, 24:Article 4279, 2024.
- [12] Shintaro Sukegawa, Atsushi Fujimura, Atsushi Taguchi, et al. Identification of osteoporosis using ensemble deep learning model with panoramic radiographs and clinical covariates. *Scientific Reports*, 12:6088, 2022.
- [13] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.