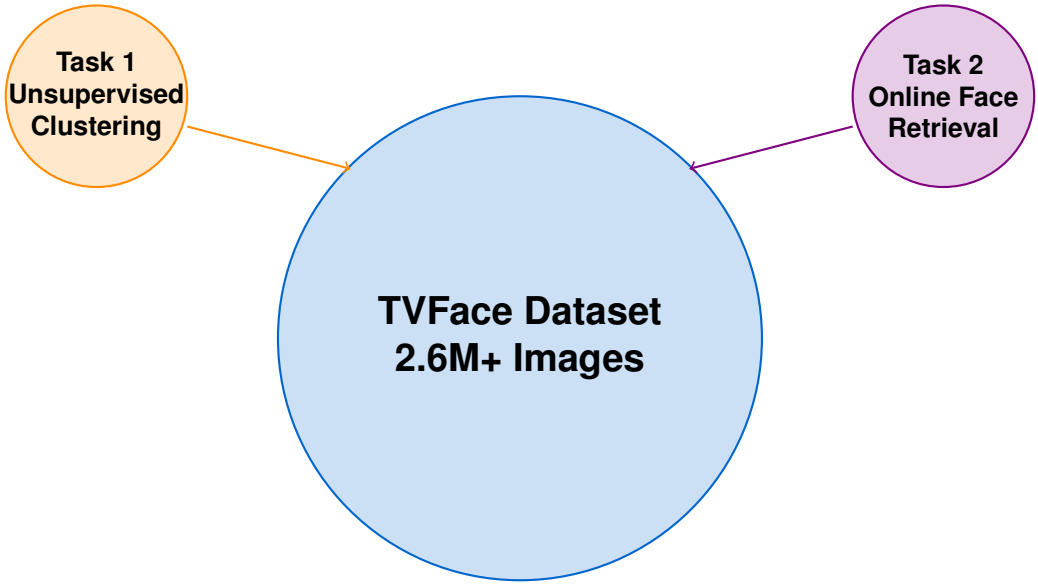
TVFace Challenge 2026

Large-Scale Facial Clustering & Retrieval Competition



Unprecedented Demographic Diversity

Competition Timeline:

Registration Opens:	March 25, 2026
Data Release:	March 26, 2026
Submission Deadline:	August 01, 2026
Results Announcement:	August 15, 2026

Organized by: Machvis
Contact: tvface-challenge@cvrc.org

Contents

1 Executive Summary

Competition Overview: The TVFace Challenge 2026 introduces two groundbreaking tasks in face recognition: **Unsupervised Facial Clustering** and **Online Face Retrieval**. With over 2.6 million images spanning diverse demographics, this competition challenges participants to develop scalable, fair, and robust algorithms for real-world face analysis applications.

The TVFace Challenge 2026 represents a paradigm shift in face recognition research, moving beyond traditional supervised learning to address the practical challenges of organizing and retrieving faces in massive, unlabeled datasets. This competition focuses on two critical tasks that reflect real-world deployment scenarios.

1.1 Key Innovation Areas

- **Unsupervised Learning:** Develop clustering algorithms without identity labels
- **Scalability:** Handle millions of images efficiently
- **Online Processing:** Real-time face retrieval systems
- **Demographic Fairness:** Ensure equitable performance across all groups
- **Practical Deployment:** Solutions ready for real-world applications

1.2 Expected Breakthroughs

- Novel unsupervised clustering algorithms for large-scale face data
- Efficient online retrieval systems with sub-second response times
- Bias-aware algorithms that work fairly across demographics
- Scalable solutions for industrial-scale face organization
- Reproducible benchmarks for future research

2 Dataset Description

2.1 TVFace Dataset Overview

The TVFace dataset comprises **2,609,210 face images** extracted from television media, representing the largest demographically diverse face dataset available for research. The dataset captures real-world variations in pose, lighting, expression, and aging effects.

Dataset Highlights:

- **Scale:** 2,609,210 total images
- **Identities:** 28,955 of unique individuals
- **Diversity:** Comprehensive demographic coverage
- **Quality:** High-resolution images with rich metadata
- **Annotations:** Age, gender, ethnicity, and quality scores

2.2 Dataset Visual Overview

Figure 1: TVFace Dataset Identity Collage: Multiple images of the same individual (Identity 3450) showing variations in age, pose, expression, lighting conditions, and temporal changes across different broadcast appearances

Identity Collage Highlights: The above collage demonstrates the rich temporal and appearance variations within a single identity cluster. This showcases the core challenge of the competition - clustering faces of the same person across different ages, poses, expressions, and broadcast conditions.

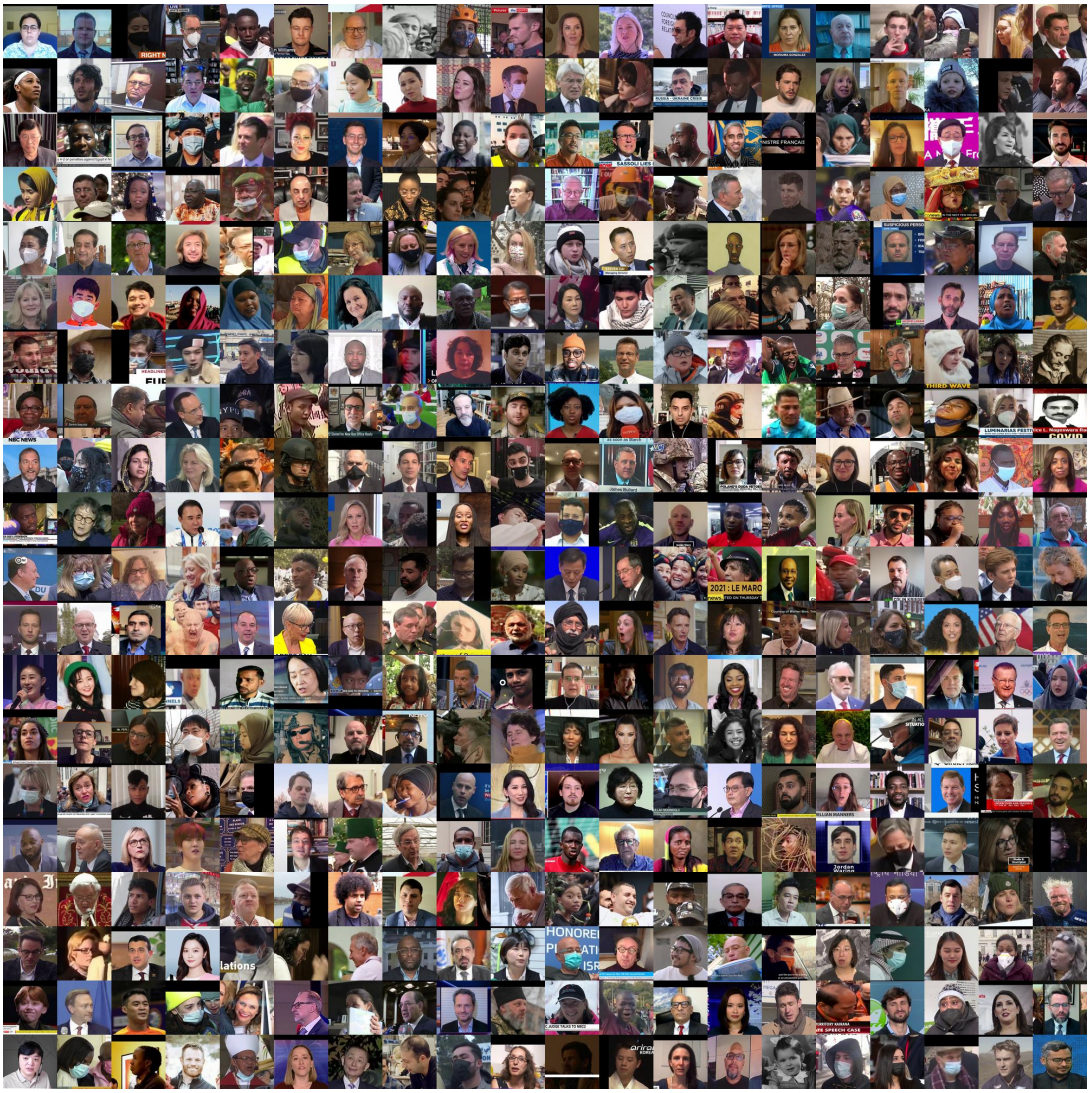


Figure 2: TVFace Dataset Identity Collage: Multiple images of the various individual showing variations in demography, age, pose, expression, lighting conditions, and temporal changes across different broadcast appearances

2.3 Identity Cluster Size Distribution

Table 1: Identity Cluster Size Distribution (Actual Data)

Identity Size Range	**Number of Identities**	**Percentage**	**Description**
10-24 images	14,477	50.0%	Small clusters (median)
25-99 images	11,582	40.0%	Medium clusters
100-999 images	2,607	9.0%	Large clusters
1000+ images	289	1.0%	Celebrity clusters
Total	**28,955**	**100%**	**Long-tail distribution**

Long-tail Characteristics: The dataset exhibits a strong long-tail distribution where 50% of identities have only 10-24 images (around the median of 24), while the largest celebrity cluster contains 21,983 images. This creates significant clustering challenges.

2.4 Expression Distribution Analysis

Expression	**Image Count**	**Percentage**	**Avg Confidence**
Neutral	1,048,995	39.39%	0.790
Sad	464,547	17.44%	0.684
Happy	464,049	17.42%	0.867
Angry	363,260	13.64%	0.717
Fear	268,530	10.08%	0.674
Surprise	47,920	1.8%	0.716
Disgust	6,072	0.23%	0.640

Table 2: Expression Distribution in TVFace Dataset (2.6M+ Images Analysis)

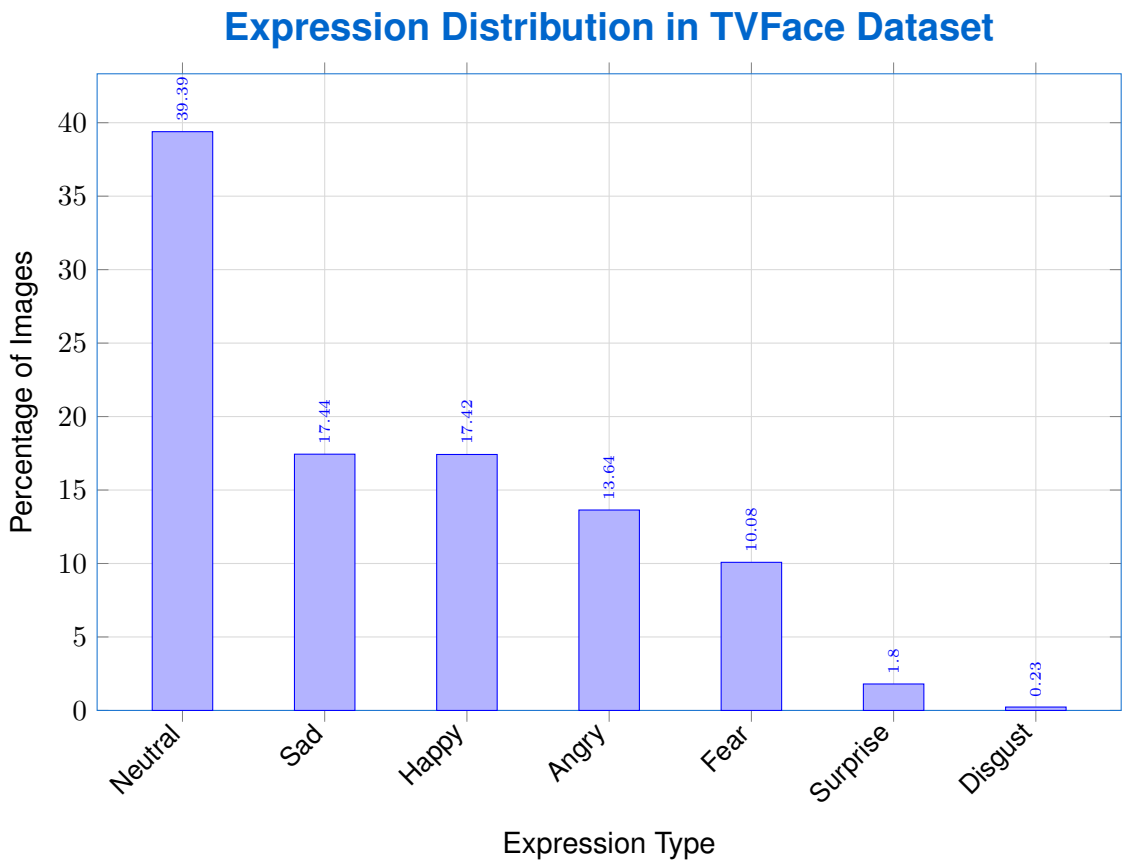


Figure 3: Beautiful expression distribution showing dominant patterns in TV broadcast data

Expression Analysis Summary:

- **Total Images Analyzed:** 2,663,373
- **Most Common Expression:** Neutral (39.4%)
- **Least Common Expression:** Disgust (0.2%)
- **Highest Confidence:** 0.867
- **TV Broadcast Context:** Professional media with controlled lighting

2.5 Head Pose Distribution

Table 3: Head Pose Statistics (Actual Analysis)

Pose Component	Mean (degrees)	Std Dev (degrees)	TV Characteristics
Yaw (Left-Right)	0.43	23.47	Wide variation, all angles
Pitch (Up-Down)	-4.27	9.25	Slight downward (camera angle)
Roll (Tilt)	-0.997	7.98	Nearly level, small variations

Pose Analysis Insights:

- **Yaw Variation:** ±23.47° indicates full profile coverage (left and right views)
- **Pitch Bias:** -4.27° shows typical TV camera angles (slightly looking down)
- **Roll Stability:** ±7.98° demonstrates mostly upright head positions
- **Broadcasting Context:** Reflects professional TV production standards

2.6 Dataset Complexity Visualization

Table 4: Dataset Complexity Metrics

Complexity Factor	Measurement	Impact on Clustering
Scale	2,609,210 images	Memory and computation limits
Identity Count	28,955 unique people	Large search space
Size Variance	10 to 21,983 per identity	Imbalanced clusters
Pose Range	Yaw: ±70°, Pitch: ±30°	View-invariant challenges
Expression Variety	7 categories with probabilities	Appearance variation
Temporal Span	Multi-year collection	Aging effects

Key Clustering Challenges:

- **Long-tail Distribution:** 50% of identities have 24 images, but max = 21,983
- **Temporal Variations:** Same person across multiple years of broadcasts
- **Pose Diversity:** Full 360° coverage with professional TV angles
- **Expression Context:** TV-specific expression patterns (neutral dominant)
- **Scale Requirements:** Algorithms must handle 2.6M+ images efficiently

2.7 Demographic Distribution Analysis

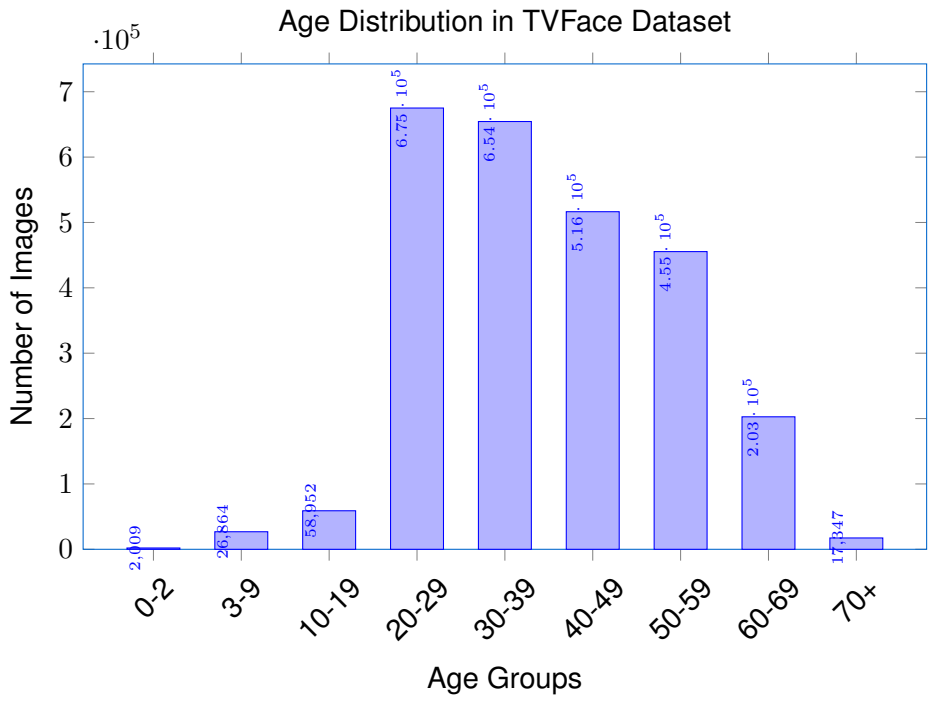


Figure 4: Age distribution showing concentration in working-age demographics

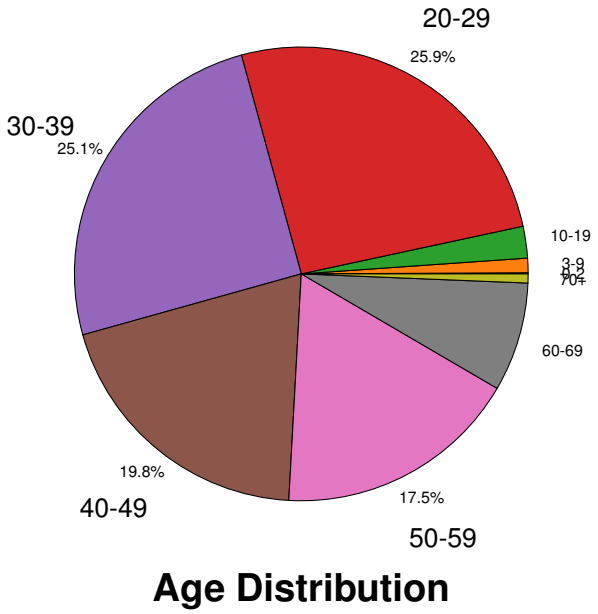


Figure 5: Pie chart showing age distribution with dominant working-age groups

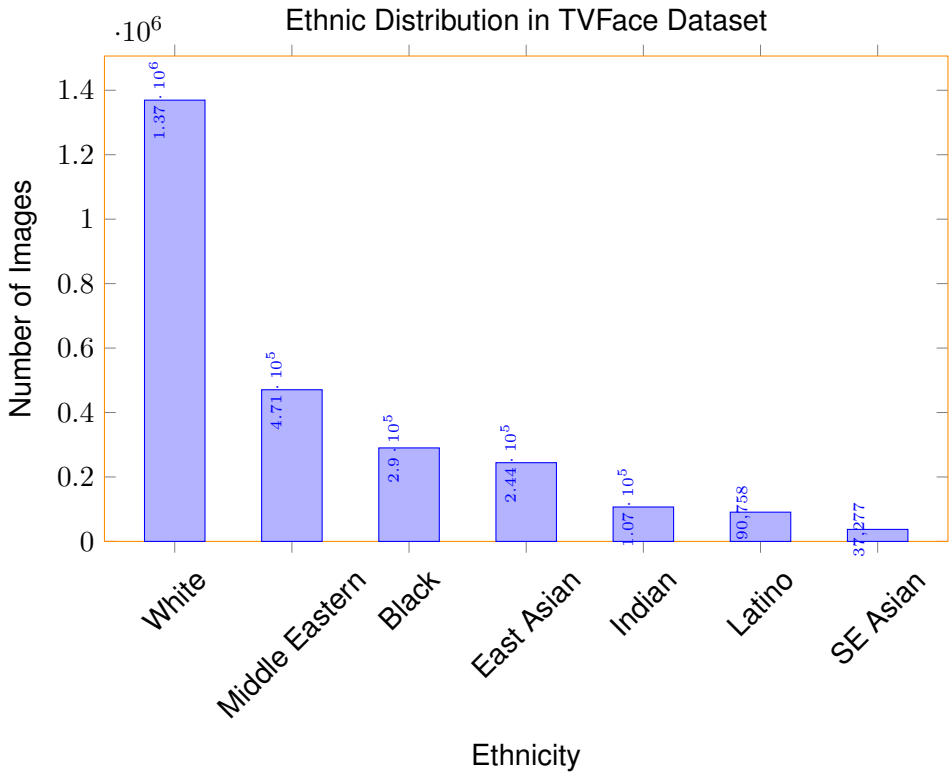


Figure 6: Ethnic distribution showing global representation

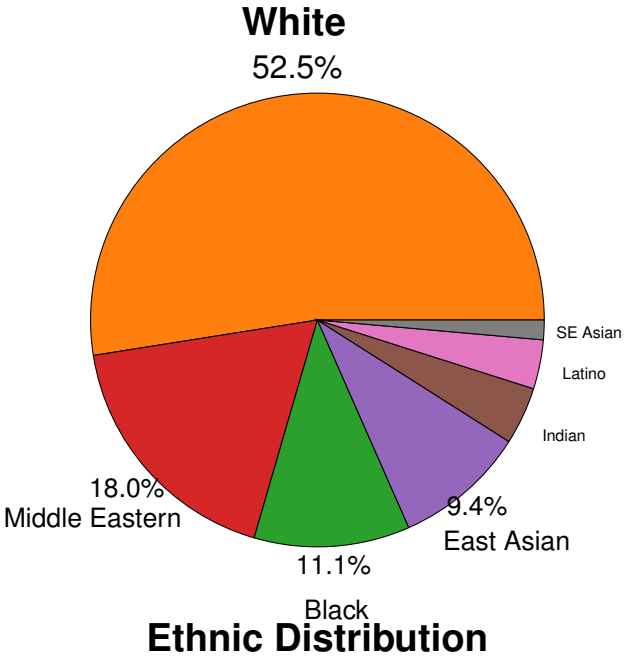


Figure 7: Pie chart showing ethnic diversity with White and Middle Eastern as dominant groups

2.8 Gender Distribution

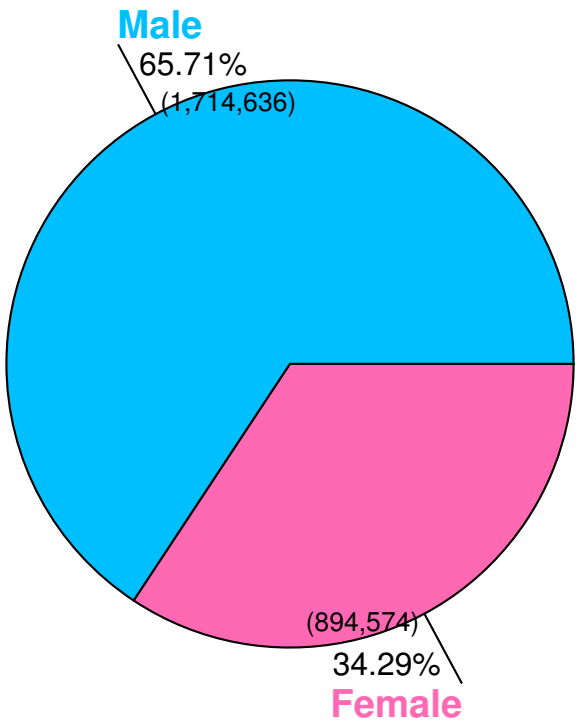


Figure 8: Corrected pie chart showing gender distribution in the TVFace dataset

Table 5: Complete Demographic Statistics Summary

Category	Subcategory	Count	Percentage
Gender	Male	1,714,636	65.71%
	Female	894,574	34.29%
Ethnicity	White	1,369,323	52.49%
	Middle Eastern	470,658	18.04%
	Black	290,143	11.12%
	East Asian	244,298	9.36%
	Indian	106,753	4.09%
	Latino Hispanic	90,758	3.48%
	Southeast Asian	37,277	1.43%
Age	20-29	675,048	25.88%
	30-39	654,423	25.08%
	40-49	516,456	19.79%
	50-59	455,416	17.46%
	60-69	202,695	7.77%
	10-19	58,952	2.26%
	70+	17,347	0.66%
	3-9	26,864	1.03%
	0-2	2,009	0.08%

3 Competition Tasks

3.1 Task 1: Unsupervised Facial Clustering

Challenge: Organize 2.6M+ face images into identity clusters without any labeled training data.

3.1.1 Task Description

Participants must develop unsupervised algorithms that can:

- Automatically discover identity clusters in unlabeled face data
- Handle varying cluster sizes (from single images to thousands)
- Maintain clustering quality across demographic groups
- Scale efficiently to millions of images
- Provide meaningful cluster representations

3.1.2 Key Challenges

- **Scale:** Processing 2.6M+ images efficiently
- **Diversity:** Handling demographic variations
- **Aging:** Clustering faces across different ages
- **Quality:** Dealing with varying image quality
- **Imbalance:** Managing clusters of different sizes

3.1.3 Evaluation Methodology

- **Clustering Purity:** Fraction of correctly clustered faces
- **Normalized Mutual Information (NMI):** Information overlap measure
- **Adjusted Rand Index (ARI):** Similarity to ground truth clustering
- **Demographic Fairness:** Performance consistency across groups
- **Scalability:** Computational efficiency metrics

3.1.4 Technical Specifications

- **Input:** Raw face images (no identity labels)
- **Output:** Cluster assignments for each image
- **Constraints:** Memory usage < 64GB, Runtime < 48 hours
- **Hardware:** Standard GPU configurations (V100/A100)

3.2 Task 2: Online Face Retrieval

Challenge: Build real-time face retrieval systems that can find similar faces instantly from massive databases.

3.2.1 Task Description

Participants must create online systems that can:

- Index millions of face images for fast retrieval
- Process query faces in real-time (< 1 second)
- Return relevant matches ranked by similarity
- Handle continuous database updates
- Maintain performance under concurrent queries

3.2.2 System Requirements

- **Latency:** Query response time < 1 second
- **Throughput:** Handle 100+ concurrent queries
- **Accuracy:** High precision in top-k results
- **Scalability:** Support dynamic database growth
- **Robustness:** Consistent performance across demographics

3.2.3 Evaluation Metrics

- **Precision@K:** Accuracy of top-k retrieved results
- **Recall@K:** Coverage of relevant results
- **Mean Average Precision (mAP):** Overall retrieval quality
- **Query Latency:** Average response time
- **Throughput:** Queries per second capacity

3.2.4 Technical Architecture

- **Indexing:** Efficient face embedding indexing
- **Search:** Approximate nearest neighbor search
- **Ranking:** Similarity-based result ranking
- **Caching:** Smart caching for popular queries
- **Load Balancing:** Distributed query processing

4 Dataset Structure and Format

4.1 Directory Organization

```
tvface_dataset/
          by_label/
                0/
                      abcnews_frame_20211215180237107908_face_0.jpg
                      abcnews_frame_20211215180237107908_face_1.jpg
                      ... (all images for identity 0)
                1/
                      bbcnews_frame_20220103142156789012_face_0.jpg
                      ... (all images for identity 1)
                2/
                      ... (all images for identity 2)
                ... (folders for each identity label)
          annotations/
                annotations.json            # Main annotation file
                demographic_summary.csv     # Aggregated demographic stats
                clustering_ground_truth.csv # For evaluation only
          query/
                easy/
                      query_001 . jpg
                      query_002 . jpg
                      ... ( other query images for easy retrieval task )
                medium/
                      query_001 . jpg
                      query_002 . jpg
                      ... ( other query images for medium retrieval task )
                difficult/
                      query_001 . jpg
                      query_002 . jpg
                      ... ( other query images for difficult retrieval task )
          splits/
                clustering_images.txt
                retrieval_database.txt
                query_images.txt
          baselines/
                clustering_baseline.py
                retrieval_baseline.py
          evaluation_scripts/
```

4.2 Data Format Specifications

4.2.1 Image Organization

Images are organized by identity labels in the `by_label/` directory. Each subdirectory contains all images belonging to a specific identity. The directory structure follows:

- **Identity Folders:** Each folder is named with the identity label (0, 1, 2, ...)
- **Image Naming:** Images retain their original descriptive names
- **Format:** All images are in JPEG format with RGB color space
- **Resolution:** 224x224 pixels, aligned and cropped faces
- **Quality:** 95% JPEG quality with minimal compression artifacts

4.2.2 Annotation Format (annotations.json)

The main annotation file contains detailed information for each image in JSON format:


```
{
  "labels": {
    "abcnews_frame_20211215180237107908_face_0": {
      "label": 83,
      "mask": 0.0,
      "attributes": {
        "age": {
          "0-2": 0.0,
          "3-9": 0.0,
          "10-19": 0.0,
          "20-29": 0.0,
          "30-39": 0.0,
          "40-49": 0.01,
          "50-59": 0.18,
          "60-69": 0.68,
          "70+": 0.13
        },
        "gender": {
          "Male": 1.0,
          "Female": 0.0
        },
        "race": {
          "White": 0.38,
          "Black": 0.0,
          "Latino Hispanic": 0.0,
          "East Asian": 0.0,
          "Southeast Asian": 0.0,
          "Indian": 0.0,
          "Middle Eastern": 0.61
        },
        "expression": {
          "angry": 0.08,
          "disgust": 0.01,
          "fear": 0.14,
          "happy": 0.0,
          "sad": 0.63,
          "surprise": 0.0,
          "neutral": 0.13
        },
        "pose": {
          "yaw": -3.26,
          "pitch": -4.65,
          "roll": -0.83
        }
      }
    }
  }
}
```

4.2.3 Annotation Field Descriptions

Table 6: Annotation Schema for Individual Images

Field	Description	Type
Root Level		
image_id	Unique image identifier (filename without extension)	String
label	Identity cluster label assigned to the image	Integer
mask	Binary mask indicating if image should be excluded (0.0 = include)	Float
Age Attributes		
0-2, 3-9, 10-19, etc.	Probability distribution over age groups	Float [0.0-1.0]
Gender Attributes		
Male	Probability of being male	Float [0.0-1.0]
Female	Probability of being female	Float [0.0-1.0]
Race Attributes		
White	Probability of White ethnicity	Float [0.0-1.0]
Black	Probability of Black ethnicity	Float [0.0-1.0]
Latino Hispanic	Probability of Latino/Hispanic ethnicity	Float [0.0-1.0]
East Asian	Probability of East Asian ethnicity	Float [0.0-1.0]
Southeast Asian	Probability of Southeast Asian ethnicity	Float [0.0-1.0]
Indian	Probability of Indian ethnicity	Float [0.0-1.0]
Middle Eastern	Probability of Middle Eastern ethnicity	Float [0.0-1.0]
Expression Attributes		
angry	Probability of angry expression	Float [0.0-1.0]
disgust	Probability of disgust expression	Float [0.0-1.0]
fear	Probability of fear expression	Float [0.0-1.0]
happy	Probability of happy expression	Float [0.0-1.0]
sad	Probability of sad expression	Float [0.0-1.0]
surprise	Probability of surprise expression	Float [0.0-1.0]
neutral	Probability of neutral expression	Float [0.0-1.0]
Pose Attributes		
yaw	Head rotation around vertical axis (degrees)	Float
pitch	Head rotation around horizontal axis (degrees)	Float
roll	Head rotation around depth axis (degrees)	Float

4.2.4 Key Annotation Properties

- **Probabilistic Labels:** All demographic and expression attributes are provided as probability distributions, reflecting annotation uncertainty
- **Multi-label Support:** Each image can have probabilities across multiple categories (e.g., mixed ethnicity)
- **Pose Estimation:** Continuous values for head pose in degrees
- **Quality Control:** Mask field allows filtering of low-quality or problematic images
- **Identity Clustering:** Label field provides ground truth for clustering evaluation

4.2.5 Usage Guidelines

Important Notes for Participants:

- For Task 1 (Clustering): Identity labels are withheld during competition
- Demographic attributes can be used for bias analysis and fairness evaluation
- Probability distributions should be interpreted as confidence scores
- Images with mask=1.0 should be excluded from analysis
- Pose values can help with data augmentation and robustness testing

5 Submission Requirements

5.1 Submission Format Overview

To maintain a fair and transparent evaluation process, all participants must follow the standardized submission formats and resource constraints outlined below. Submissions will be validated automatically using an open-source evaluation script.

5.2 Track A: Unsupervised Face Clustering

5.2.1 Required Submission File

File: `submission_clusters.json`

Listing 1: Clustering Submission Format

```
{
    "skynews_000102394.jpg": "cluster_001",
    "cnn_000203848.jpg": "cluster_001",
    "dw_000020102.jpg": "cluster_002",
    ...
}
```

Format Requirements:

- Each image filename maps to a cluster ID
- Cluster IDs can be arbitrary strings (e.g., "C123", "cluster_001")
- Cluster IDs must be consistent and non-empty
- Maximum cluster count: **50,000**

5.3 Track B: Face Retrieval

5.3.1 Required Submission File

File: `submission_retrieval.json`

Listing 2: Retrieval Submission Format

```
{
    "query_0001.jpg": [
        "skynews_000102394.jpg",
        "cnn_000203848.jpg",
        ...
    ],
    "query_0002.jpg": [
        "dw_000203945.jpg",
        ...
    ]
}
```

Format Requirements:

- Each query image filename maps to a ranked list of retrieved images
- Maximum of **100 results per query**
- Results must be ranked in **descending similarity order**
- All retrieved filenames must exist in the dataset

5.4 Final Submission Package

Participants must submit a single ZIP file with the following structure:

Listing 3: Required Directory Structure

```
tvface26_submission/
          metadata.txt
          task1/
                submission_clusters.json
          task2/
                submission_retrieval.json
          optional/
              log.txt / clusters_summary.csv
```

Critical Requirements:

- File naming must **not be changed**
- Submit ZIP file as: `teamname_tvface26.zip`
- Include `metadata.txt` with team information
- ZIP file size limit: **500 MB**

5.5 Local Evaluation

The official evaluation script is open-sourced for local testing: **TVFace-26 Evaluation Toolkit**

5.5.1 Evaluation Commands

```
# Evaluate clustering task
python eval_cluster.py \
    --submission submission_clusters.json \
    --groundtruth annotations.json

# Evaluate retrieval task
python eval_retrieval.py \
    --submission submission_retrieval.json \
    --groundtruth retrieval_gt.json
```

5.6 Resource & Submission Constraints

Table 7: Submission Constraints and Limits

Constraint	Limit
Max Cluster Count	50,000 (for Task 1)
Max Retrieval Results/Query	100 (for Task 2)
Submission ZIP Size	500 MB
Allowed Runtime	30 minutes per task on evaluation server
Allowed Memory	8 GB RAM, 2 CPU cores (GPU optional)
Max Submissions per Day	2 (per team, per track)
Final Leaderboard Freeze	One week before challenge deadline

5.7 Submission Portal

Evaluation Platform: All final submissions must be uploaded to the **TVFace-26 evaluation platform** powered by EvalAI.
Key Features:

- Results appear on public leaderboard immediately after evaluation
- Test set submissions remain hidden until leaderboard freeze
- Automatic validation and scoring
- Real-time feedback and error reporting

Platform Status: **Coming Soon** - Will be available with dataset release

5.8 Submission Validation Checklist

Table 8: Pre-Submission Validation Checklist

darkgreen!20 **Validation Item**	**Status**
JSON format is valid and parseable	☐
lightgray!30 All required files are included	☐
File naming convention followed exactly	☐
lightgray!30 Resource constraints satisfied	☐
Local evaluation script passes	☐
lightgray!30 ZIP file size under 500 MB	☐
Metadata.txt contains team information	☐

Pro Tips for Successful Submission:

1. Test your submission with the official evaluation script **before** uploading
2. Use the daily submission limit strategically for iterative improvements
3. Monitor the public leaderboard for baseline comparisons
4. Submit early to avoid last-minute technical issues
5. Keep backup copies of all submission files

5.9 Detailed Schedule

Table 9: Competition Schedule 2026

Date	Milestone	Description
March 15, 2026	Challenge Announcement	Official CFP posted on website
March 25, 2026	Dataset & Baselines Released	Public release of train/val sets, evaluation toolkit
April 10, 2026	Evaluation Server Opens	EvalAI server goes live with dev leaderboard
May 1, 2026	Submission Phase Begins	Participants can submit results (dev phase)
May 10, 2026	Challenge Webinar #1	Introduction, walkthrough, and Q&A session
June 15, 2026	Leaderboard Update	First snapshot of top submissions (optional blog post)
July 1, 2026	Challenge Webinar #2	Technical deep-dive & expert panel
August 1, 2026	Submission Deadline (Test)	Final date to submit results on hidden test set
August 2, 2026	Leaderboard Freeze	Public leaderboard frozen for final ranking
August 10, 2026	Report Submission Due	Short system description paper (max 4 pages)
August 20, 2026	Winner Notification	Top teams notified, final checks, preprint paper prep
September 10, 2026	Joint Competition Paper	Organizers + top teams submit a joint paper
December 13, 2026	Results Announced	Winners presented and honored during official track

Key Timeline Highlights:

- **5-month competition period:** From dataset release (March 25) to final submission (August 1)
- **Two educational webinars:** May 10 (introduction) and July 1 (technical deep-dive)
- **Development phase:** 3 months of active submission and leaderboard feedback
- **Final evaluation:** Test set submissions with 1-day freeze period
- **Conference presentation:** Results announced at major venue in December

Important Deadlines:

- **August 1, 2026:** Final submission deadline - no extensions
- **August 10, 2026:** System description papers due
- **September 10, 2026:** Joint paper submission with organizers

6 Rules and Guidelines

6.1 Participation Rules

- **Team Size:** Maximum 5 members per team
- **Eligibility:** Open to academic and industry participants
- **Registration:** One account per team
- **Multiple Submissions:** Up to 5 submissions per day
- **Final Selection:** Teams select best submission before deadline

6.2 Technical Requirements

- **External Data:** Allowed if publicly available and declared
- **Pre-trained Models:** Permitted with proper attribution
- **Code Sharing:** Public sharing encouraged after competition
- **Reproducibility:** All results must be reproducible
- **Hardware Limits:** Standard GPU configurations only

6.3 Ethical Guidelines

Important: All participants must adhere to ethical AI principles:

- Respect privacy and data protection regulations
- Consider fairness and bias implications
- Avoid discriminatory applications
- Follow institutional ethics protocols
- Report potential misuse scenarios

7 Frequently Asked Questions

8 Contact Information

Competition Organizers

- **General Inquiries:** tvface-challenge@cvrc.org
- **Technical Support:** support@tvface-challenge.org
- **Data Access:** data@tvface-challenge.org
- **Media Relations:** media@cvrc.org

8.1 Organizing Committee

9 Conclusion

The TVFace Challenge 2026 represents a significant milestone in face recognition research, introducing two critical tasks that address real-world deployment challenges. By focusing on unsupervised clustering and online retrieval, this competition pushes the boundaries of what's possible in large-scale face analysis.

We invite researchers, practitioners, and students to participate in this groundbreaking challenge. Together, we can advance the field of computer vision while ensuring that face recognition technology serves all members of society fairly and effectively.

Join us in shaping the future of face recognition technology!