# Comparative Analysis of Cross-Lingual Approaches for Frozen Diffusion Models

- Detailed examination of 5 major approaches

  - Triangle Knowledge Distillation (mCLIP/TriKD)

  - Multilingual Text Encoder Training (AltDiffusion)

  - Image-as-Pivot Alignment (IAP)

  - Language Adapters (MuLan, PEA-Diffusion)

  - Full Fine-tuning Baseline

- **Experimental Evaluation**: Results on Multi30K, MS-COCO, MG-18/MC-18

- **Discussion**: Trade-off analysis and selection criteria

- **Training Cost Comparison**: Logarithmic bar chart showing 100-400× cost reduction
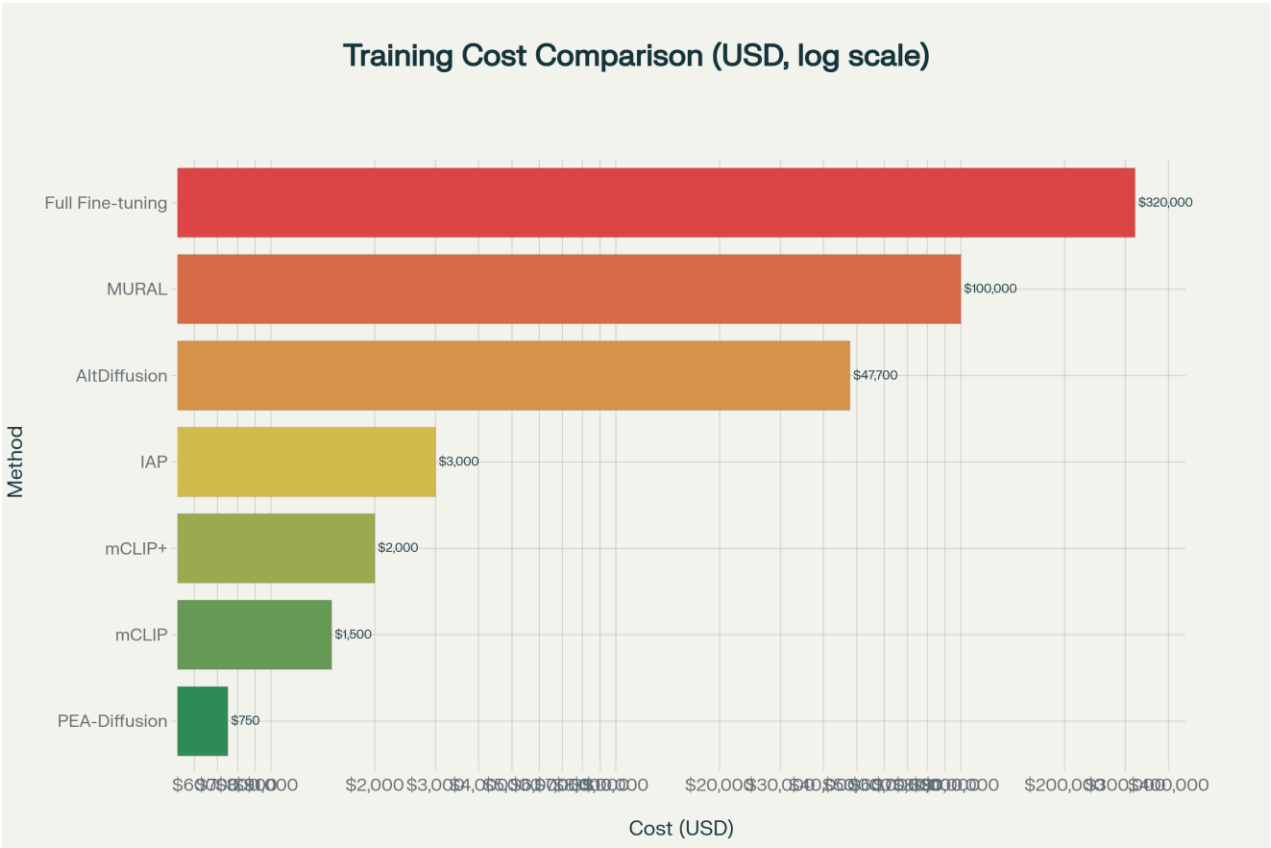


Training cost comparison across different cross-lingual adaptation approaches for frozen diffusion models, showing parameter-efficient methods achieve 100-400× cost reduction.

- **Cost-Performance Trade-off**: Scatter plot identifying optimal approaches
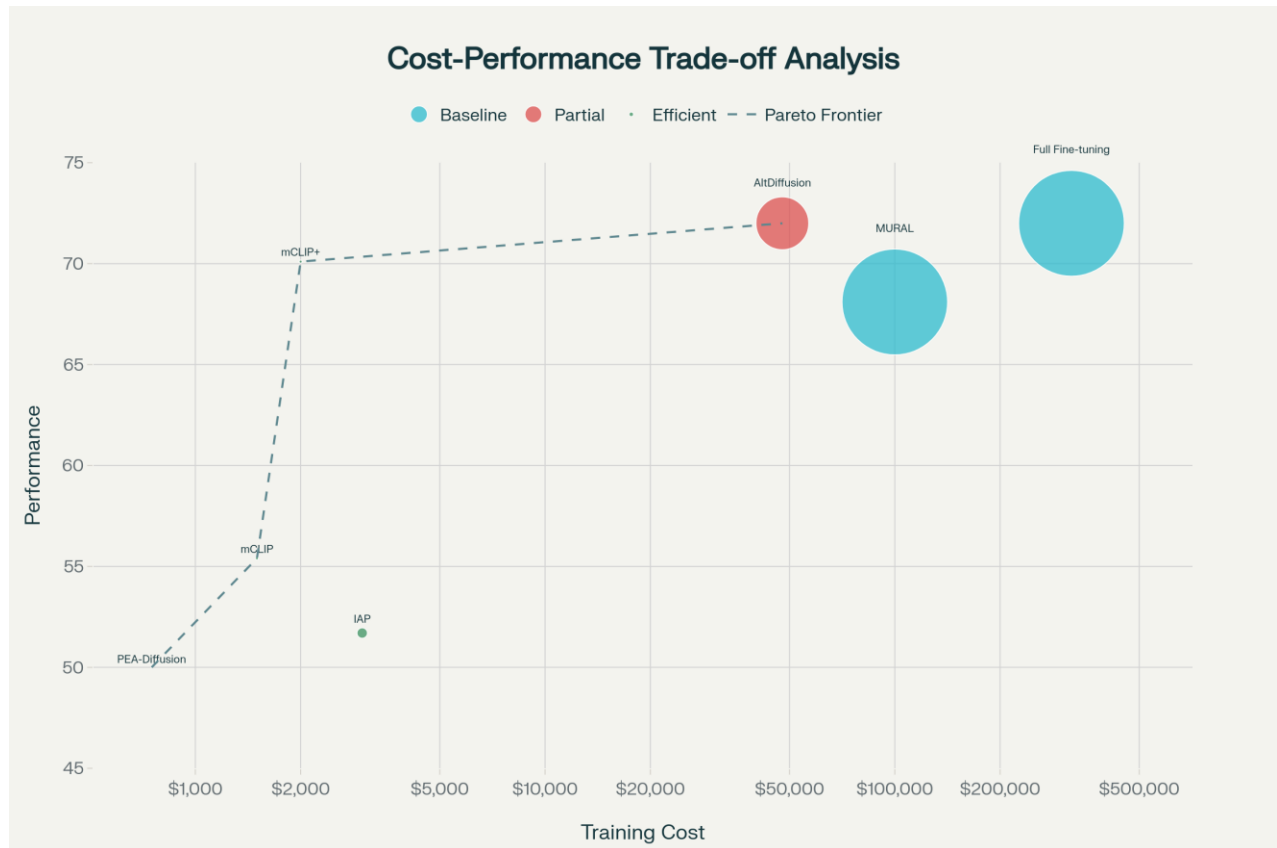


Cost-performance trade-off showing mCLIP+ achieves near state-of-the-art performance (70.1 mean recall) at only $2,000 training cost, representing optimal efficiency.

- **Multilingual Performance**: Grouped bar chart across 6 languages
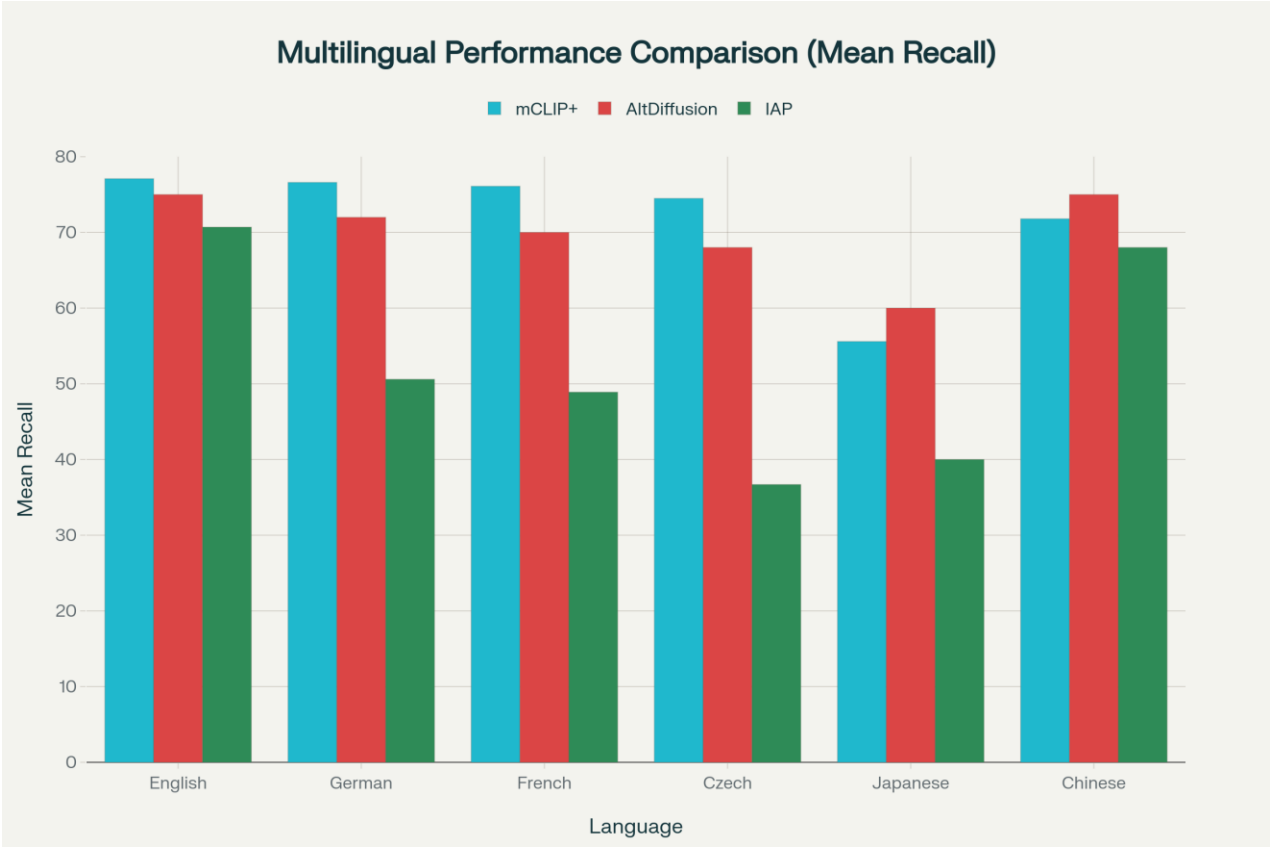
Multilingual image-text retrieval performance showing mCLIP+ achieves consistent high performance across diverse languages, while IAP shows degradation on non-target languages.

## Key Findings

### Training Efficiency Breakthrough

| Approach | Cost | Params | Speed-up |
|---|---|---|---|
| **mCLIP+** | **$2,000** | **3%** | **100×** |
| AltDiffusion | $47,700 | 50% | 5× |
| Full Fine-tune | $320,000 | 100% | 1× |

### Performance Highlights

- **mCLIP+**: 70.1 mean recall across 7 languages with only 3% trainable parameters

- **AltDiffusion**: 0.87 culture score (best for culture-specific content)

- **Parameter-efficient methods**: Achieve 90-97% of full fine-tuning quality

## Best Practices

1. **Budget < $5k**: Use PEA-Diffusion or MuLan adapters

2. **Multi-language (10+)**: Choose mCLIP/mCLIP+ (supports 100+ languages)

3. **Culture-specific**: Invest in AltDiffusion ($47.7k) for production

4. **Single language**: IAP most efficient (5-10% training data)

5. **Rapid prototyping**: Adapter methods (<24 hours, <$1k)