

Motivation:

- LinkedIn/corporate headshots require specific quality standards
- Manual review is time-consuming and subjective
 - No existing labeled datasets for professional portrait attributes



Project Definition:

Build a model that automatically evaluates 6 quality attributes of professional portrait photos

Models:

- Swin Transformer Tiny (28M parameters, pretrained ImageNet)
- 6 independent binary classification heads

Data:

- Source: FFHQ dataset (70,000 high-quality face images)
 - Generated: Around 6k labeled synthetic portraits via Gemini API
 - Labels: 6 binary attributes per image

- Input:
 - A single headshot image.
- Output: Vector with 6 binary attributes:
 - Lighting Even and Frontal
 - Background Clean and Non-Distracting
 - Business or Professional Attire Visible
 - Neutral Professional Facial Expression
 - Face Properly Framed and Centered
 - Image Sharpness High

Original FFHQ (top) vs Generated with ONE Bad Attribute (bottom)

ORIGINAL
ffhq_03404



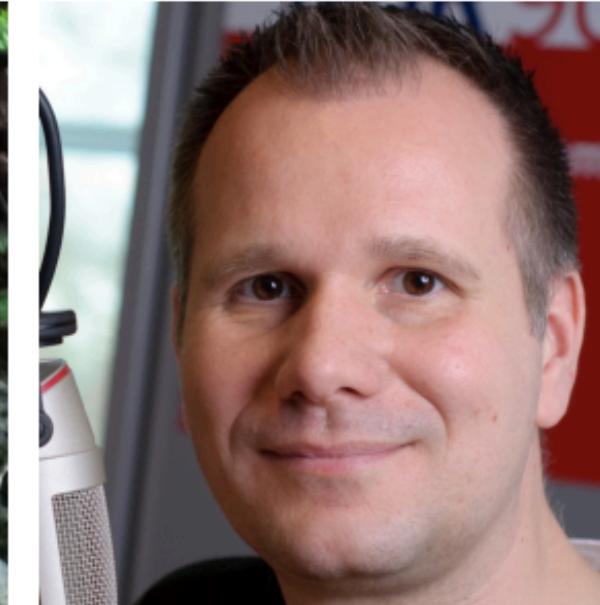
ORIGINAL
ffhq_04110



ORIGINAL
ffhq_03297



ORIGINAL
ffhq_02426



ORIGINAL
ffhq_00737



ORIGINAL
ffhq_00811



Bad Lighting
01111



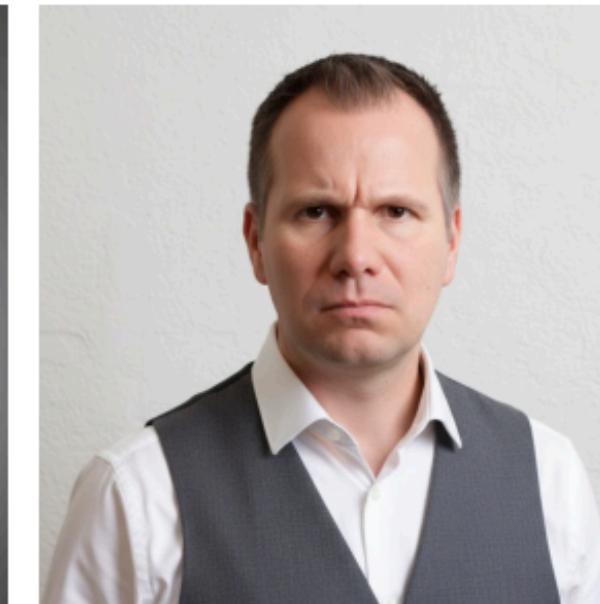
Bad Background
10111



Bad Attire
110111



Bad Expression
111011



Bad Framing
111101



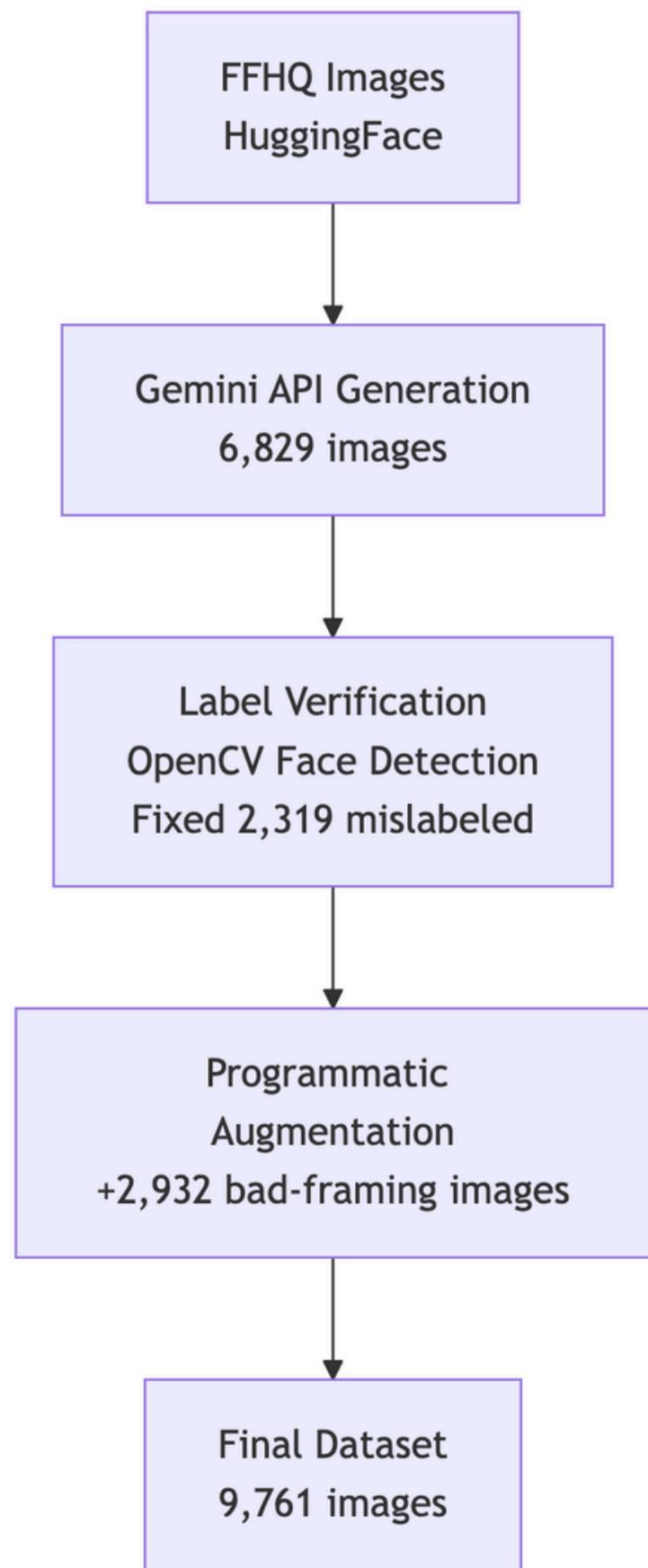
Bad Sharpness
111110



Parameter	Value
Model	Swin Transformer Tiny
Input Size	224 × 224
Batch Size	32
Epochs	10
Optimizer	AdamW (lr=5e-5)
Scheduler	OneCycleLR (max_lr=1e-4)
Loss	BCEWithLogitsLoss
Train/Val Split	80/20 (GroupShuffleSplit)

From ~15% to ~70% perfect match score

Metric	Result
Overall Accuracy	93.89%
Exact Match	69.86%
Macro F1	0.95
Training Images	9,761
Attributes Detected	6



- Successfully generated a large-scale synthetic dataset of 9,761 professional portraits
- Perfect Match From ~15% to ~70% perfect match score by label verification, (Fixed 2,319 mislabeled images (82.8% of "bad framing" were actually good), Framing Balancing, Threshold Tuning

GOOD
(Centered, proper size)



GOOD
(Centered, proper size)



GOOD
(Centered, proper size)

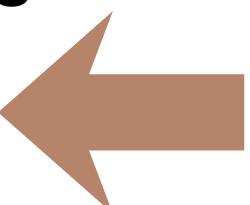
BAD
(Off-center or wrong size)



BAD
(Off-center or wrong size)



Framing GenAI examples



GOOD
(Centered, proper size)



BAD
(Off-center or wrong size)



BAD
(Off-center or wrong size)



ORIGINAL
Good Framing (label=1)

OFF_CENTER_LEFT
Face pushed right
(left side cropped)
(label=0)



EXTREME CLOSEUP
Too close - only
part of face visible
(label=0)

OFF_CENTER_RIGHT
Face pushed left
(right side cropped)
(label=0)



TOO_FAR
too small - subject
lost in frame
(label=0)

HEAD_CUTOFF
Top of head
cut off
(label=0)

EXTREME CLOSEUP
Too close - only
part of face visible
(label=0)



Golden Dataset: Model Predictions vs Human-Verified Labels

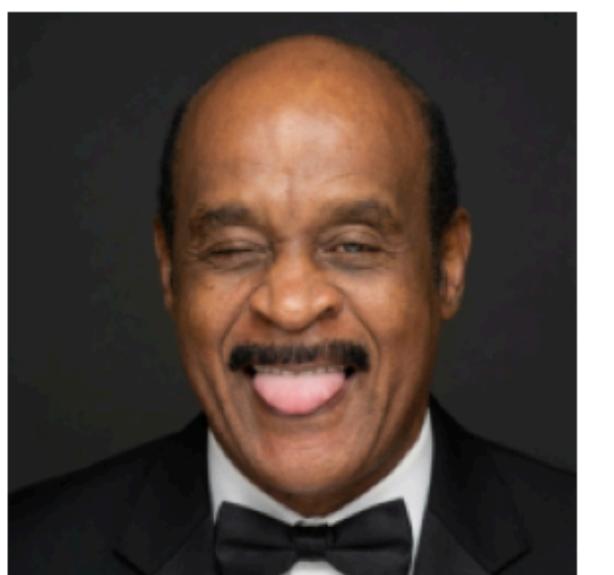
Golden Image #2
PERFECT



Golden Image #3
PERFECT



Golden Image #5
PERFECT



GROUND TRUTH vs PREDICTION

Attribute	True	Pred
Lighting Even and Frontal	1	1
Background Clean and Non-Dis	1	1
Business or Professional Att	0	0
Neutral Professional Facial	1	1
Face Properly Framed and Cen	1	1
Image Sharpness High	1	1

OVERALL ACCURACY (Hamming)
EXACT MATCH (all 6 correct)
MACRO F1 SCORE

97.0%
83.3%
0.982

Golden Image #1
HAS ERRORS



GROUND TRUTH vs PREDICTION

Attribute	True	Pred
Lighting Even and Frontal	0	0
Background Clean and Non-Dis	1	1
Business or Professional Att	0	0
Neutral Professional Facial	1	1
Face Properly Framed and Cen	0	1
Image Sharpness High	1	1

Golden Image #4
HAS ERRORS



GROUND TRUTH vs PREDICTION

Attribute	True	Pred
Lighting Even and Frontal	0	1
Background Clean and Non-Dis	1	1
Business or Professional Att	0	0
Neutral Professional Facial	1	1
Face Properly Framed and Cen	1	1
Image Sharpness High	0	0

Golden Image #9
HAS ERRORS



GROUND TRUTH vs PREDICTION

Attribute	True	Pred
Lighting Even and Frontal	1	1
Background Clean and Non-Dis	0	0
Business or Professional Att	1	1
Neutral Professional Facial	0	0
Face Properly Framed and Cen	0	1
Image Sharpness High	0	0

Impact of Our Improvements



Baseline
Face Framing: 82%
Overall Acc: 91.2%
Exact Match: 57.0%

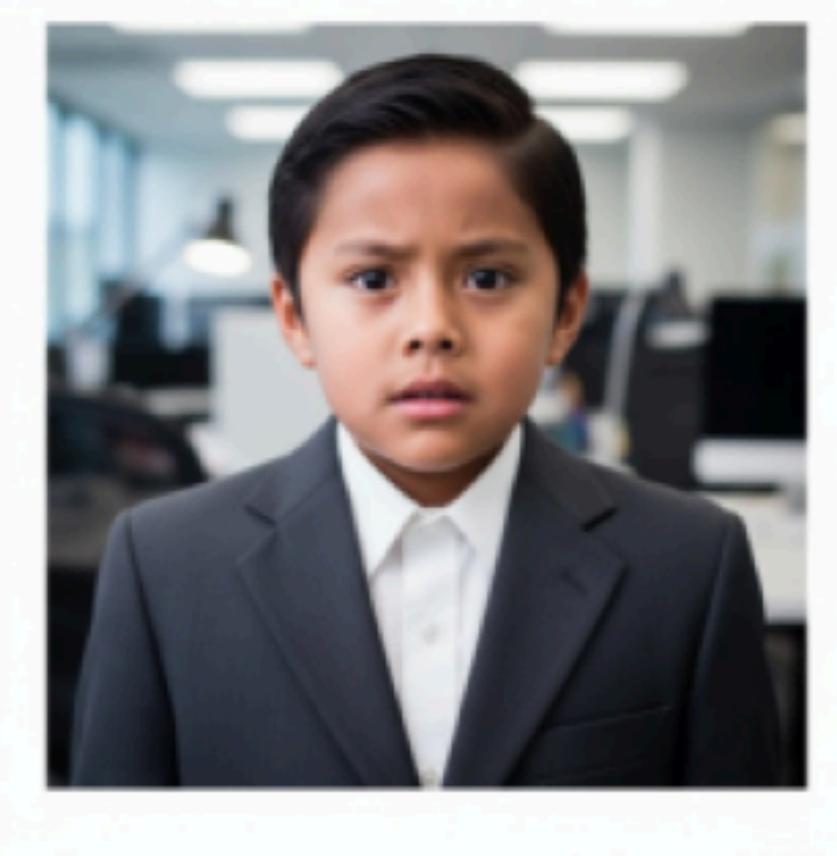


IMAGE 3: PERFECT MATCH

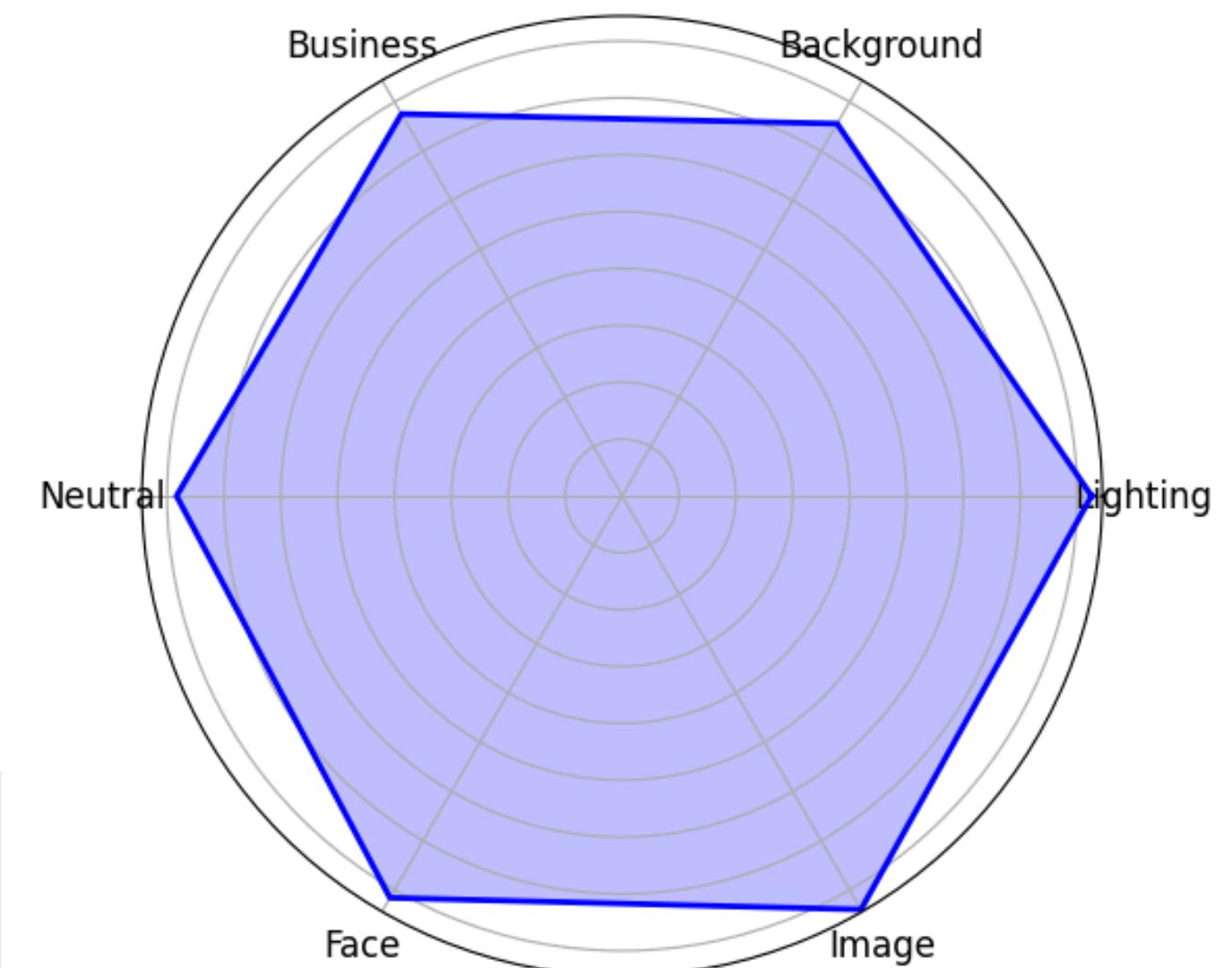
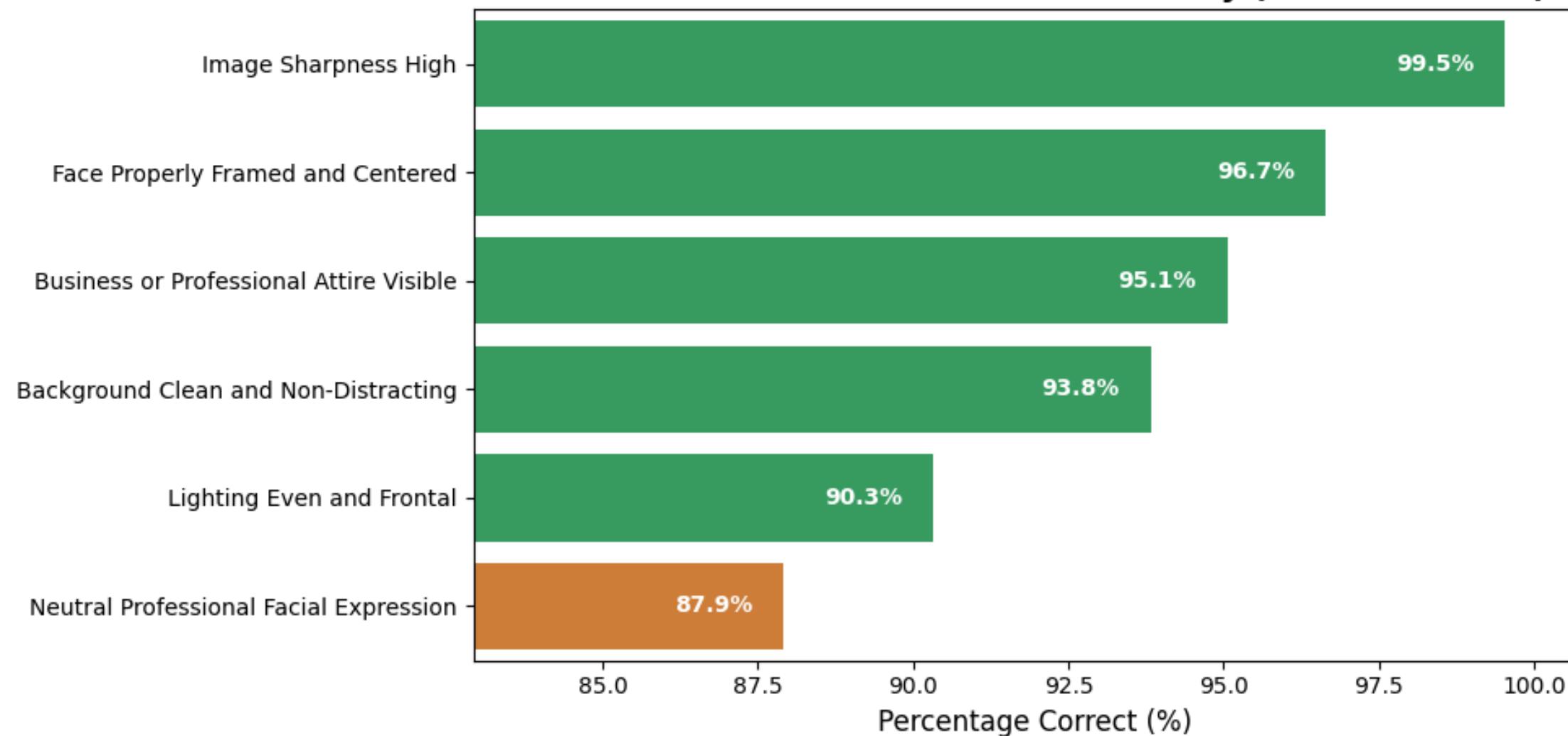
- Lighting Even and Frontal P:1 | T:1
- Background Clean and Non-Distractin P:1 | T:1
- Business or Professional Attire Vis P:1 | T:1
- Neutral Professional Facial Express P:0 | T:0
- Face Properly Framed and Centered P:1 | T:1
- Image Sharpness High P:1 | T:1



IMAGE 4: HAS ERRORS

- Lighting Even and Frontal P:1 | T:1
- Background Clean and Non-Distractin P:1 | T:1
- Business or Professional Attire Vis P:1 | T:1
- Neutral Professional Facial Express P:1 | T:1
- Face Properly Framed and Centered P:1 | T:0
- Image Sharpness High P:1 | T:1

Stage	Face Framing	Overall Acc	Exact Match
Baseline	82%	91.2%	57.0%
+ Verification	83%	91.5%	59.2%
+ Balancing	97%	93.9%	69.9%

Model Capability Profile (Macro F1: 0.80)**Performance Profile: Attribute Accuracy (Macro F1: 0.95)**

Overall Accuracy 93.89%
Perfect Image Ratio: 69.86%

What We Learned	Implication
Generators ignore negative prompts	83% of "bad framing" prompts failed
Programmatic augmentation works	Sharpness: 100%, Framing: 97%
Label verification is essential	Fixed 2,319 errors automatically
Balanced data is critical	Framing improved from 82% to 97% after balancing

Idea	Current	Expected
Programmatic lighting augmentation	92%	97%+
Expression verification (emotion model)	90%	94%+
Larger model (Swin Base or Large)	93.90%	96%+

