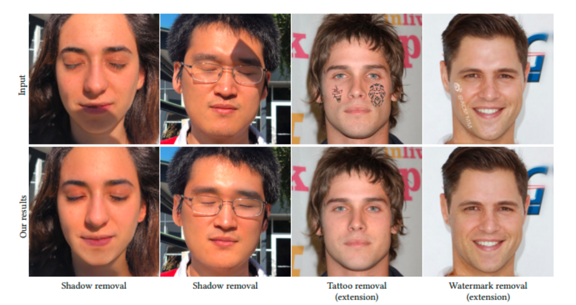
# Unsupervised Portrait Shadow Removal via Generative Priors

Yingqing He, Yazhou Xing, Tianjia Zhang, Qifeng Chen

Coby Penso

# What we'll see

Portrait images often suffer from undesirable shadows cast by casual objects or even the face itself




Portrait shadow removal as an effective, high quality, and automatic application is highly desirable.
In this paper, an Unsupervised method suggested.

# Existing Solutions



Recent state-of-the-art shadow removal and portrait shadow removal methods are based on deep learning techniques and trained on large-scale image pairs in a supervised manner.

**Better performance than traditional methods**

**Issues:**

- Preparing training data is a tedious and challenging task
- Fail on complex and various real-world images.
- Difficult training and practical usage.

# The General Idea

Viewing the shadow removal task as **Image decomposition problem**:

- Full-Shadow Image
- Shadow-Free Image
- Shadow Mask

Leverage generative priors to solve the ambiguity of image layer decomposition.

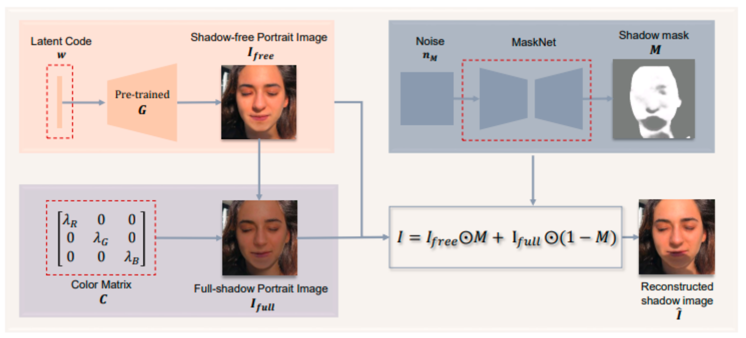
Input | Full-shadow portrait | Shadow-free portrait | Learned shadow mask | Recovered shadow portrait

Bar-Ilan University

Given and image $I$, decompose it into $I_{free}$, $I_{full}$ and $M$:

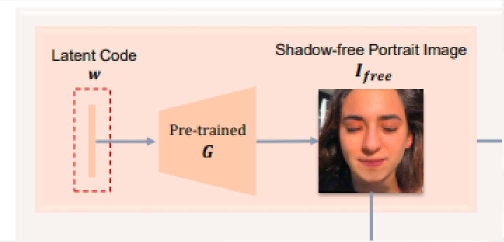
$$I = I_{free} \odot M + I_{full} \odot (1 - M)$$

Where $M$ shadow mask - denote shadow region and intensity



# Method - $I_{free}$

- Pretrained StyleGAN2, denoted $G$
- $G$ as the branch for recovering undelying $I_{free}$
- Using optimization-based GAN inversion method

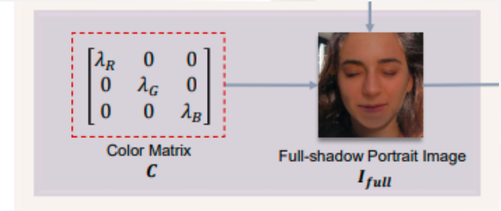


# Method - Regularize Full Shadow Image

To further regularize the full shadow image $I_{full}$, use a color matrix $C \in [0,1]^{3x3}$

$$I_{full}(x,y) = C \cdot I_{free}(x,y)$$

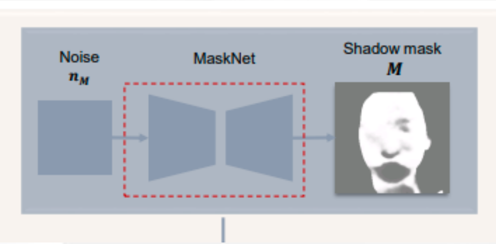
Where $(x,y)$ denote pixel positions and $\gamma$ denote learnable shadow parameters.



# Method - Shadow Mask M



- Use small network *MaskNet* $f_M$
- Take random initialized noise map $n_M$ as the input
- Reconstruct the shadow mask M




- MaskNet design as Encoder-Decoder network with skip connections (Similar to Double-DIP).
- Sigmoid at the output to further regularize $M \in (0, 1)$

# Optimization

- Optimization has to be done carefully
- How to utilize the prior knownledge of StyleGAN2
- Straight forward jointly optimizing $I_{free}$, $I_{full}$, $C$, $M$ lead to poor results

Input | Our model, No GP

- **Instead adopt Progressive optimization strategy to guide the image recovering process step by step**

# Progressive Optimization

Divide the optimization to stages

- Stage 1 - Initial face optimization
- Stage 2 - Color matrix and Shadow mask optimization
- Stage 3 - Facial features refinement

Bar-Ilan University

**Initial face optimization**

- Use GAN invertion method to project the shadowed image into StyleGAN2 latent space
- Only aim at recovering high-quality human faces through segmenting face parts with a pretrained face parsing model.
- Use LPIPS loss as the optimization goal for GAN inversion.

$$L_{LPIPS} = ||\Phi(I_{free}^{init} \odot S) - \Phi(I \odot S)||_2$$

S is the segmentation mask for the face.
Φ is the pretrained VGG-19 network (LPIPS Model)

# Stage 1 - Latent initialization

Bar-Ilan University

In order to ease the optimization process.
Found that latent initialization can help:

- Start with w that approximate the face image $I \odot S$ well.
- Randomly sample 500 latent vectors from the prior distribution.
- Generate 500 images
- Use $L_{LPIPS}$ to select best initial value for w

# Stage 1 - Summary

Bar-Ilan University

**Input:** Shadowed portrait $I$, face parsing map $S$

**Output:** Shadow-free portrait $I_{free}$, full-shadow portrait $I_{full}$, blending mask $M$

---

*Stage 1 – Initial face optimization.*

---

Sample 500 $\{z_i\}_{i=1}^{500}$ from Gaussian distribution;
Infer $w$ space latents $\{w_i\}_{i=1}^{500}$ using $\{z_i\}_{i=1}^{500}$;
Select $w_b$ which minimizes $\mathcal{L}_{LPIPS}(I, G(w_i))$;
$w_b^0 = w_b$;
**for** $k = 1$ to $K$ **do**
    $I_{free}^{init} = G(w_b^{k-1})$;
    Loss=$\mathcal{L}_{LPIPS}(I_{free}^{init}, I)$;
    Update $w_b^{k-1}$ using ADAM algorithm;
**end for**
**return** $I_{free}^{init} = G(w_b^K)$.

Bar-Ilan University

**Color matrix and shadow mask optimization**

Goals:

- Recover full-shadow portrait $I_{full}$ with color matrix C
- Estimate the shadow mask $M$ for image blending

Steps:

- Fix the latent space of StyleGAN2
- Jointly optimize color matrix C and parameters $\theta$ of MaskNet $f_M$
- Minimize the reconstruction loss

$$\hat{I}_{C,\theta} = I_{free} \cdot M + C \cdot I_{free} \cdot (1 - M)$$

$$\min_{C,\theta} ||\hat{I}_{C,\theta} - I||_2^2$$

In practice: do 50 steps of optimization Why? to guide M to learn the blending relationship instead of compensating for the face details.

# Stage 2 - Summary

---

*Stage 2 – Color matrix and shadow mask optimization.*

---

```
Randomly initialize the MaskNet $f_M^0$ and the noise map $n_M$;
$M^0 = f_M^0(n_M)$;
Initialize diagonal element of color matrix to 0.5 to obtain $C^0$;
for $p = 1$ to $P$ do
    $M^p = f_M^p(n_M)$;
    $\hat{I} = I_{free}^{init} \odot M^{p-1} + (C^{p-1} I_{free}^{init}) \odot (1 - M^{p-1})$;
    Loss=$MSE(\hat{I}, I)$;
    Update $C^{p-1}$ and $f_M^p$ using ADAM algorithm;
end for
return $M^P$ and $C^P$.
```

---

Bar-Ilan University

**Facial features refinement**
**Problem:** The face reconstruction results of the first stage may miss perceptually important face details.
**Goal:**

- Improve projection quality to StyleGAN2 to refine face details
- Using global perceptual loss to the whole face
- Also, facial feature loss to optimize face components

$$F = \{nose, eyebrow, eyeball, mouth, glasses\}$$

$$L_{feat} = \sum_{f \in F} \lambda_f \Phi(f, \hat{f})$$

$$\min_{C,w} L_{feat} + L_{LPIPS}$$

note: optimize also C, since the change of face detail may influence the estimation of full shadow images.

# Stage 3 - Summary

Bar-Ilan University

*Stage 3 – Facial features refinement.*

Initialization: $C^0 = C^P$, $w^0 = w_b^K$, $I_{free}^0 = I_{free}^{init}$;

**for** $q = 1$ to $Q$ **do**

  $I_{free}^{q-1} = G(w^{q-1})$;

  $\hat{I} = I_{free}^{q-1} \odot M^P + (C^{q-1} I_{free}^{q-1}) \odot (1 - M^P)$;

  Loss $= \mathcal{L}_{feat}(\hat{I}, I) + \mathcal{L}_{LPIPS}(\hat{I}, I)$;

  Update $C^{q-1}$ and $w^{q-1}$ using ADAM algorithm;

**end for**

$I_{free} = G(w^Q)$, $C = C^Q$, $M = M^P$, $I_{full} = C \times I_{free}$;

**return** $I_{free}$, $I_{full}$ and $M$.

# Extensions

Bar-Ilan University

Can be extended to

- Face tattoo removal
- Face watermark removal

Modify the formulation,

$$I = I_{clean} \odot M + I_{pure} \odot (1 - M)$$

with binary mask restriction, to improve performance:

$$L_{binary} = min(|M - 0|, |M - 1|)$$

# Experiments Settings

- real-world portrait shadow removal dataset [36]
- 9 subjects, 100 shadow portrait images in varies poses, shadow shapes, illumination conditions and shadow types.
- For tattoo and watermark removal, synthesize their data based on CelebA-HQ dataset
- StyleGAN2 model which is pretrained on FFHQ (256x256)

Comparing to state-of-the-art supervised shadow removal methods.

Input | ST-CGAN [30] | DAD [38] | DHAN [5] | PSM [36] | Ours | Ground truth

| | Methods | SSIM ↑ | LPIPS ↓ |
|---|---|---|---|
| Supervised methods | ST-CGAN [30] | 0.512 | 0.3031 |
| | DAD [38] | 0.603 | 0.3225 |
| | DHAN [5] | 0.629 | 0.1607 |
| | PSM [36] | **0.859** | **0.0874** |
| Unsupervised methods | DGP [24] | N/A | N/A |
| | Ours (No GP) | 0.707 | 0.3270 |
| | Ours (JO) | 0.811 | 0.1288 |
| | Ours | **0.820** | **0.1162** |

# Experiments

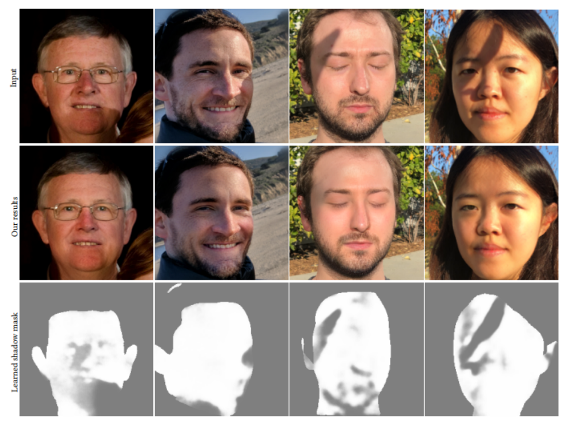
Controlled experiments

- **Joint optimization (JO):** Instead of optimizing different terms alternatively, jointly optimize the shadow-free image, color matrix and shadow mask
- **No generative priors (NO GP):** Conduct experiment without pretrained StyleGAN2 weights. Keep the same network architecture but randomly initialize the StyleGAN2 weights.

Input | Our model, No GP | Our model, JO | Our model | Ground truth

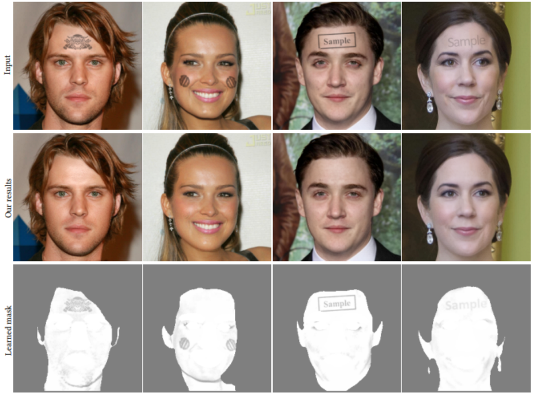
# Experiments

Observing the learned shadow mask M



Testing the method on extensions

- Tattoos
- Watermarks




Tattoo removal | Tattoo removal | Watermark removal | Watermark removal

Questions?

# Bibliography


- Unsupervised Portrait Shadow Removal via Generative Priors
https://arxiv.org/pdf/2108.03466.pdf
- Analyzing and Improving the Image Quality of StyleGAN
https://arxiv.org/pdf/1912.04958.pdf
- Shadow Removal via Shadow Image Decomposition
https://arxiv.org/abs/1908.08628