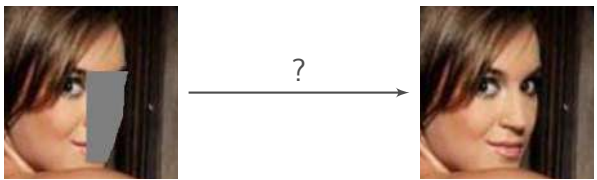
3DFaceFill

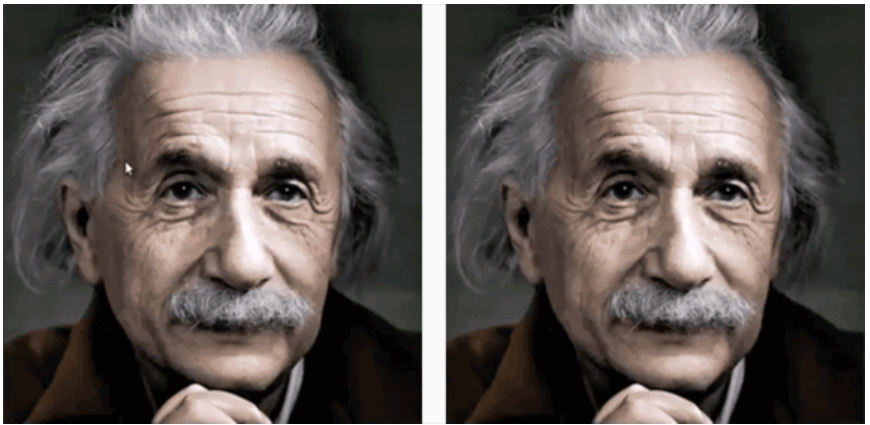
An Analysis-By-Synthesis Approach to Face Completion



Rahul Dey

MICHIGAN STATE
UNIVERSITY

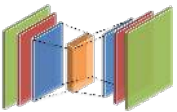
What is face completion



Source: nVidia

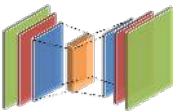
Existing approaches work in 2D

Existing approaches work in 2D



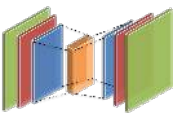
Auto encoder

Existing approaches work in 2D

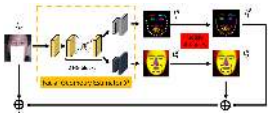


Auto encoder

Existing approaches work in 2D

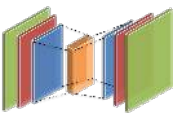


Auto encoder

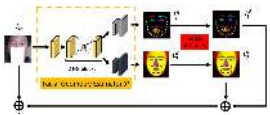


Face specific losses
genfacecompletion

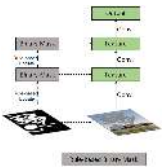
Existing approaches work in 2D



Auto encoder

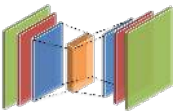


Face specific losses
genfacecompletion

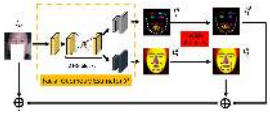


Partial Convolutions
partialconv

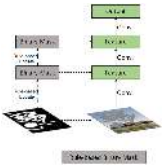
Existing approaches work in 2D



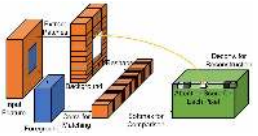
Auto encoder



Face specific losses
genfacecompletion

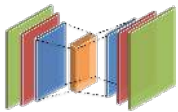


Partial Convolutions
partialconv

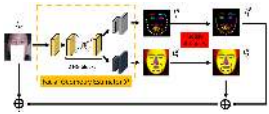


Attention
gencontextualattn

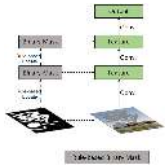
Existing approaches work in 2D



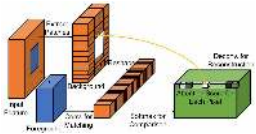
Auto encoder



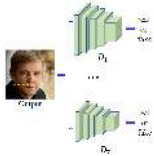
Face specific losses
genfacecompletion



Partial Convolutions
partialconv

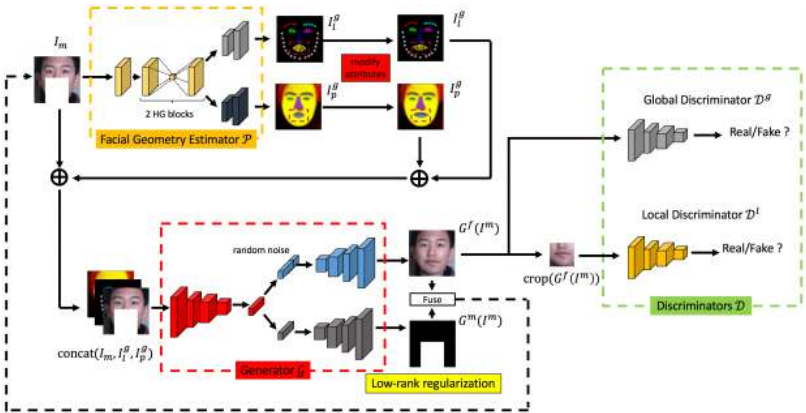


Attention
gencontextualattn

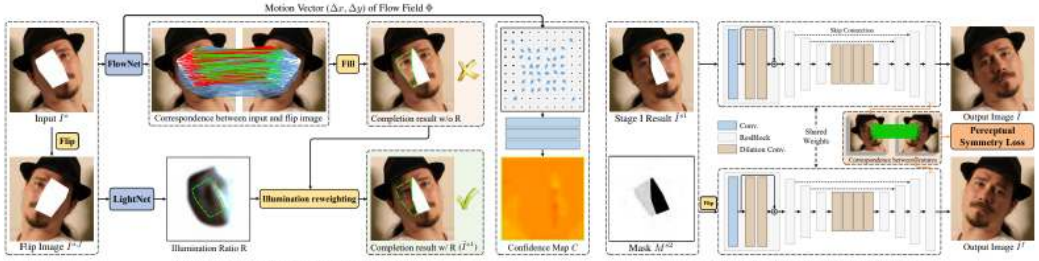


Multiple discriminators
oracleattention

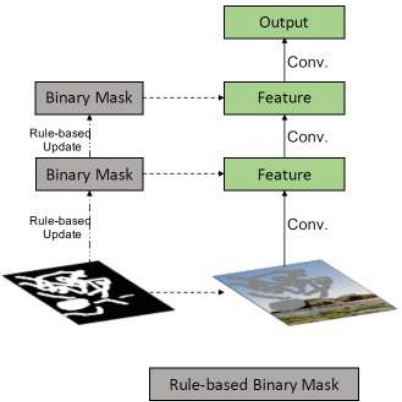
GFC (CVPR, 2017)



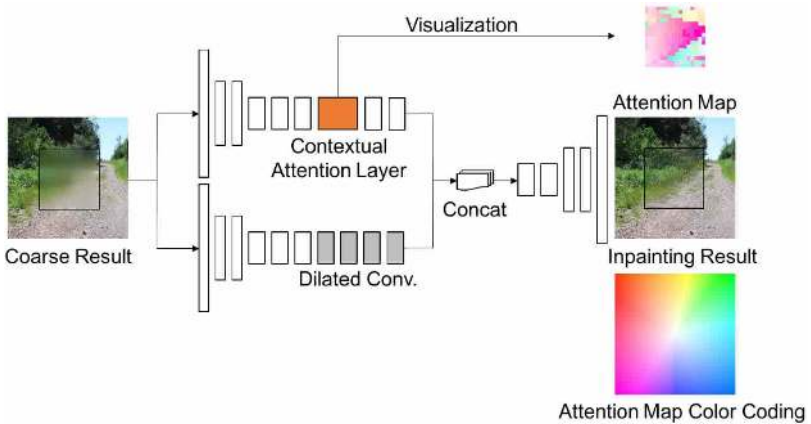
SymmFCNet (TIP, 2020)



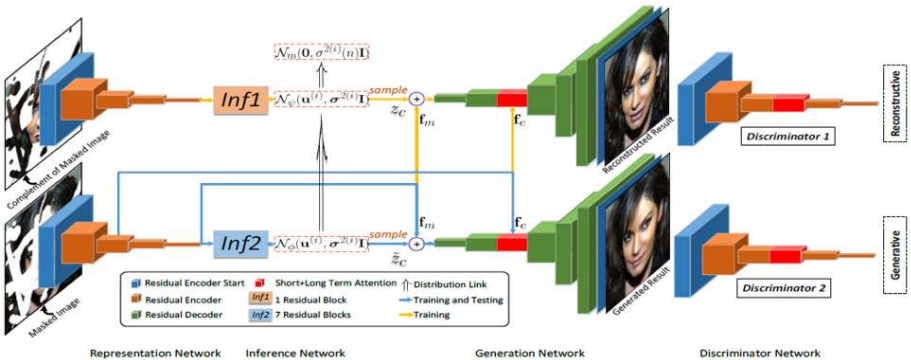
PartialConv, Nvidia (ECCV, 2018)



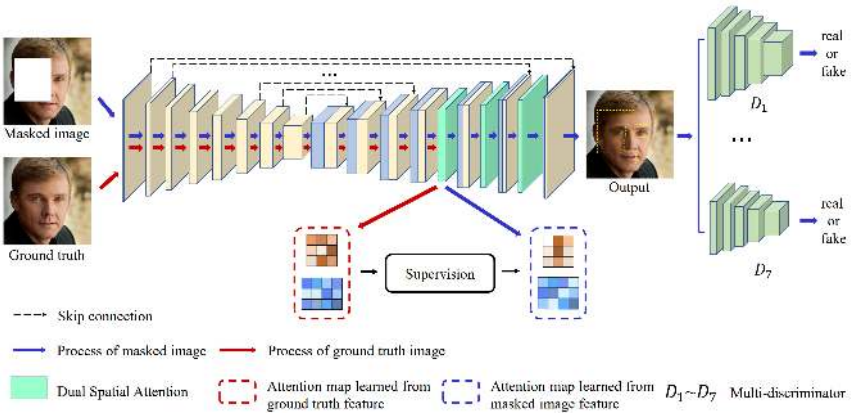
DeepFill, Adobe (ICCV, 2019)



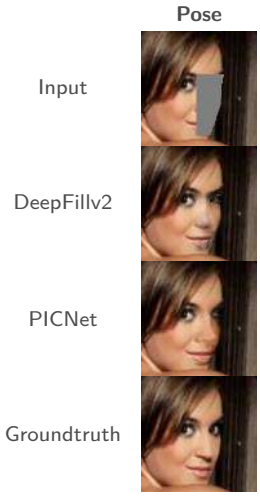
PICNet (CVPR, 2019)



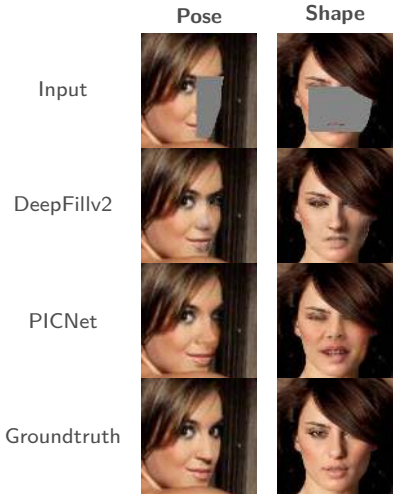
DSA (CVPR, 2020)



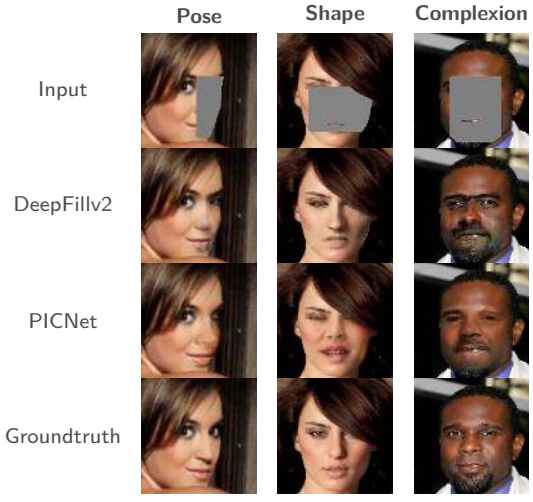
2D based FC models may fail when



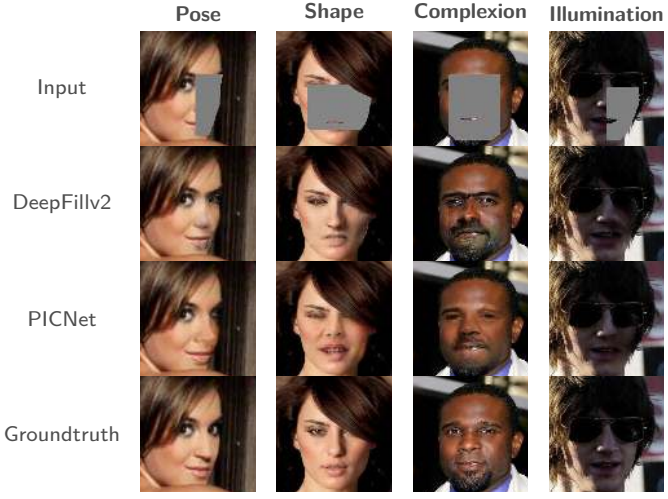
2D based FC models may fail when



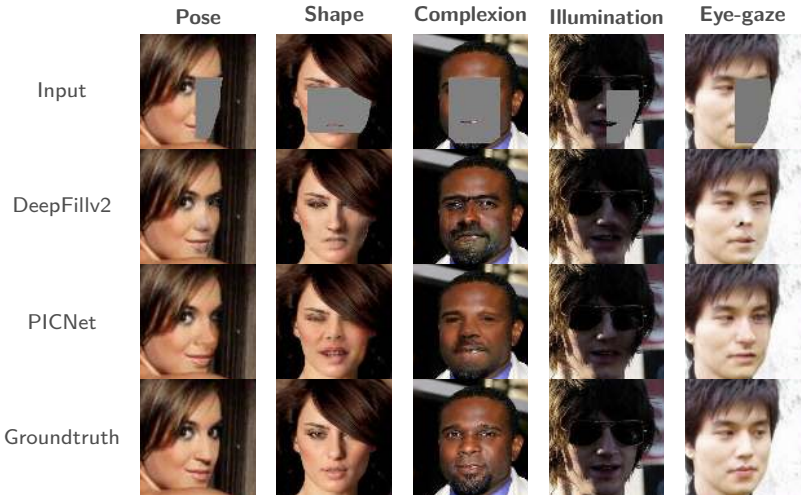
2D based FC models may fail when



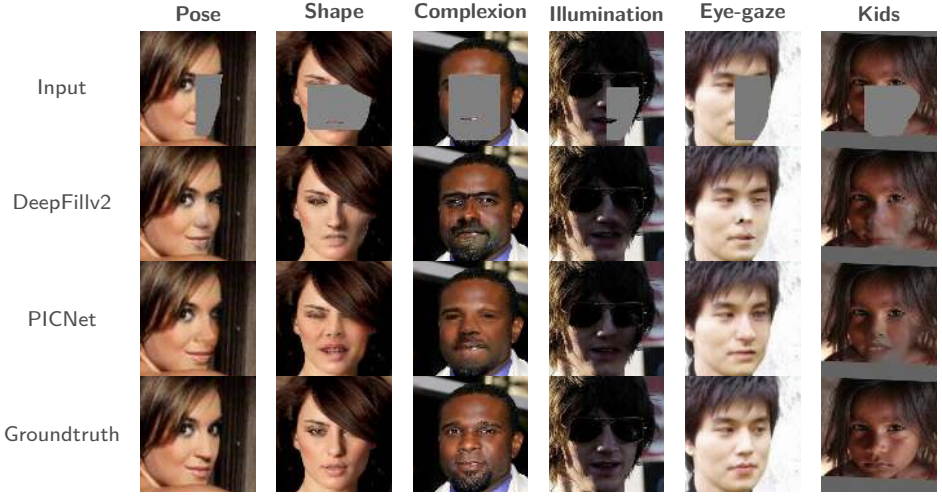
2D based FC models may fail when



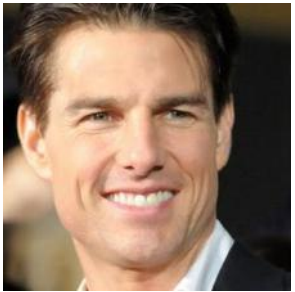
2D based FC models may fail when



2D based FC models may fail when

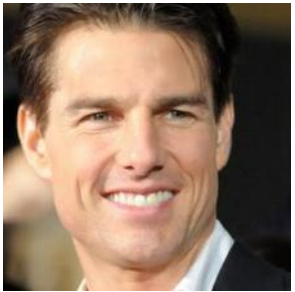


Face is a 3D Object



3D Object

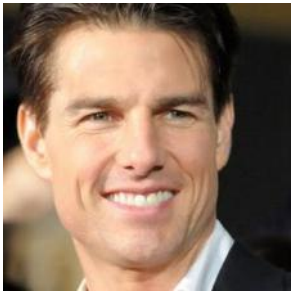
Face is a 3D Object



3D Object

Shape

Face is a 3D Object

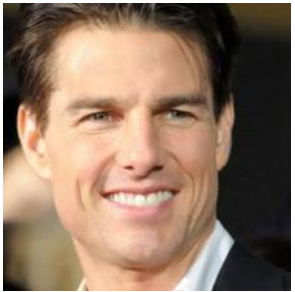


3D Object

Shape

$Pose = (s, \mathbf{R}_{3D}, \mathbf{t}_{3D})$

Face is a 3D Object



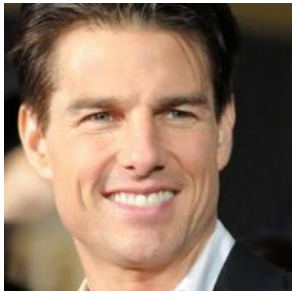
3D Object

Shape

$Pose = (s, \mathbf{R}_{3D}, \mathbf{t}_{3D})$

Albedo

Face is a 3D Object



3D Object

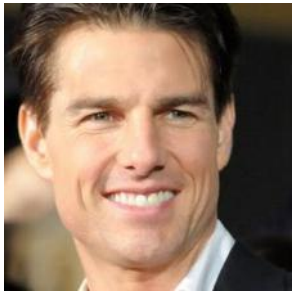
Shape

$Pose = (s, \mathbf{R}_{3D}, \mathbf{t}_{3D})$

Symmetry

Albedo

Face is a 3D Object



3D Object

Shape

$Pose = (s, \mathbf{R}_{3D}, \mathbf{t}_{3D})$

Illumination

Albedo

Symmetry

Our hypothesis

Explicit 3D
modeling of

Our hypothesis

Explicit 3D
modeling of

Shape

Our hypothesis

Explicit 3D
modeling of

Shape, Pose

Our hypothesis

Explicit 3D
modeling of

Shape, Pose, Appearance

Our hypothesis

Explicit 3D
modeling of

Shape, Pose, Appearance and Illumination

Our hypothesis

Explicit 3D
modeling of

Shape, Pose, Appearance and Illumination

using vanilla architectures

Our hypothesis

Explicit 3D modeling of >>

Shape, Pose, Appearance and Illumination

using vanilla architectures

Our hypothesis

Explicit 3D
modeling of

>>

Existing approaches

Implicit 3D
modeling of

Shape, Pose, Appearance and Illumination

using vanilla architectures

Our hypothesis		Existing approaches
Explicit 3D modeling of	>>	Implicit 3D modeling of

Shape, Pose, Appearance and Illumination

using vanilla architectures	using advanced architectures

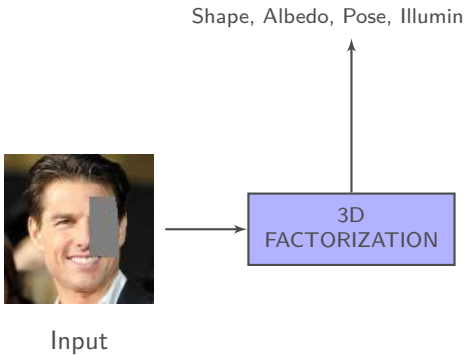
Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$

Input

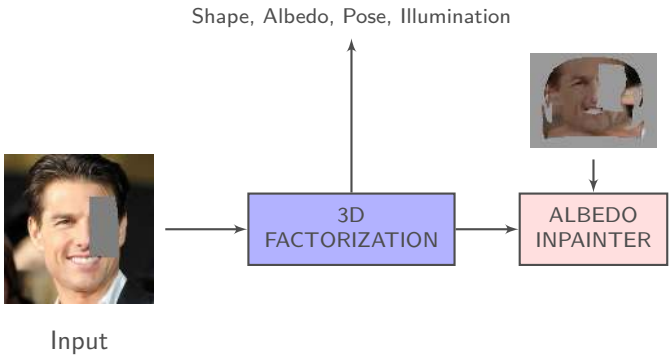
Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$



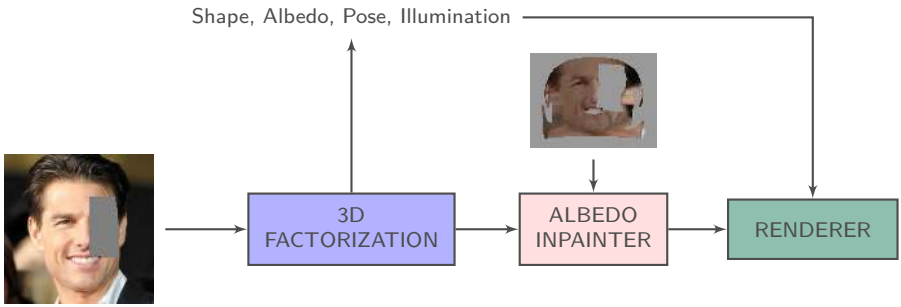
Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$



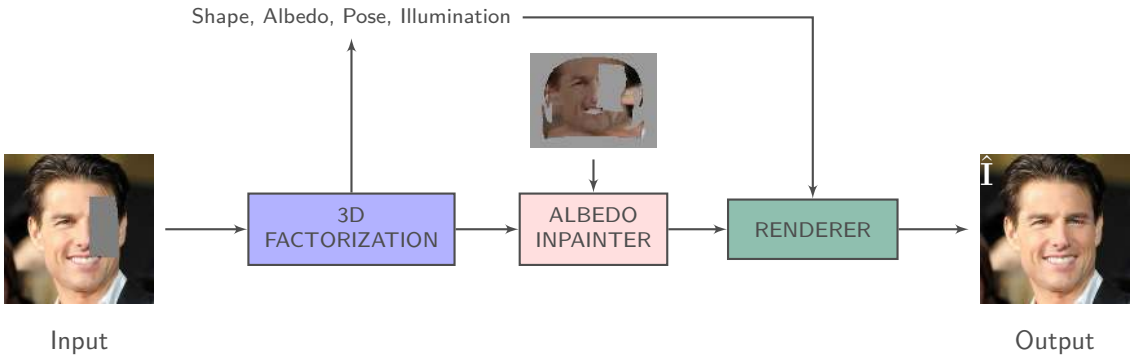
Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$



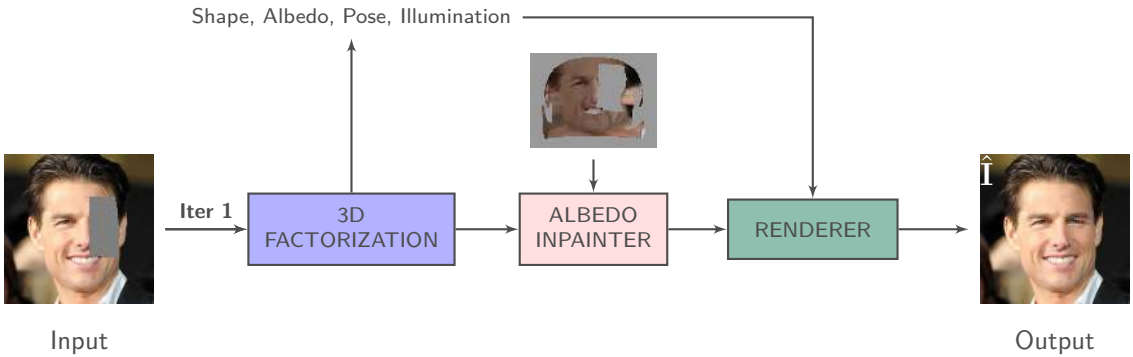
Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$



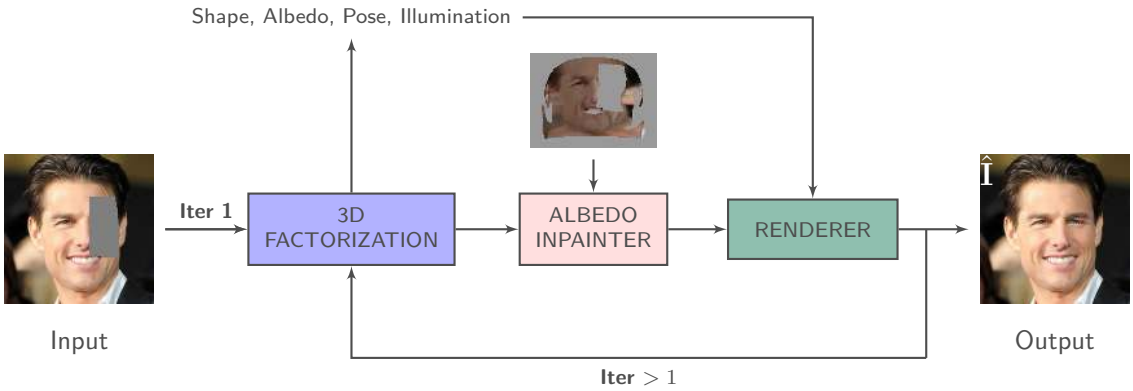
Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$

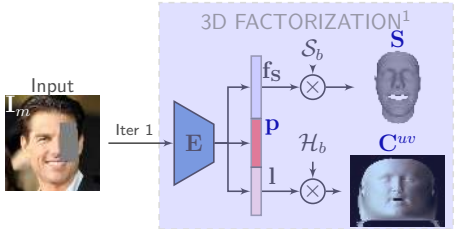


Proposed approach: 3D factorization

$$\Phi : \mathbf{I} \rightarrow (\mathbf{S}hape, \mathbf{A}lbedo, \mathbf{P}ose, \mathbf{Il}lumination) \quad (1)$$

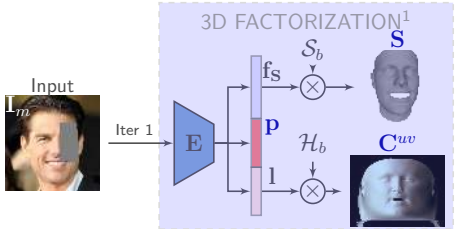


Architecture



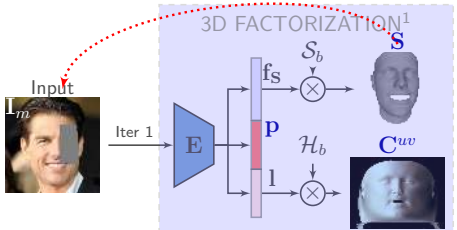
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



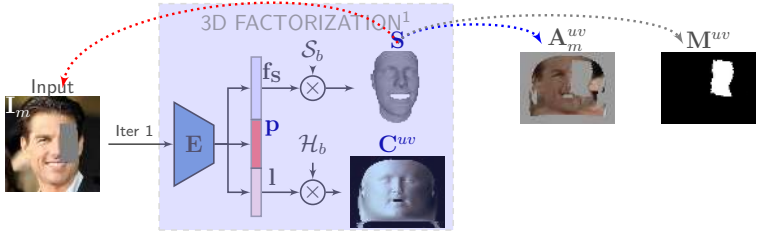
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



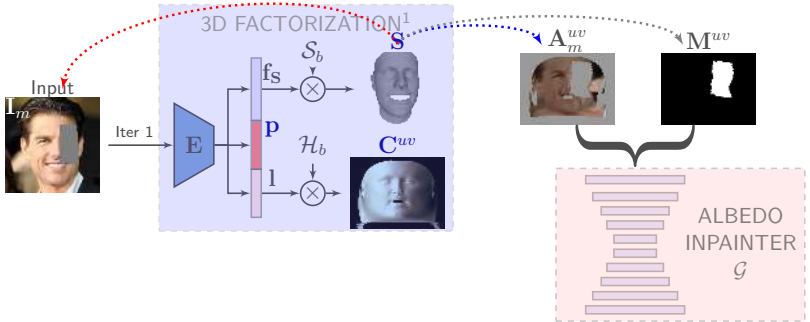
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



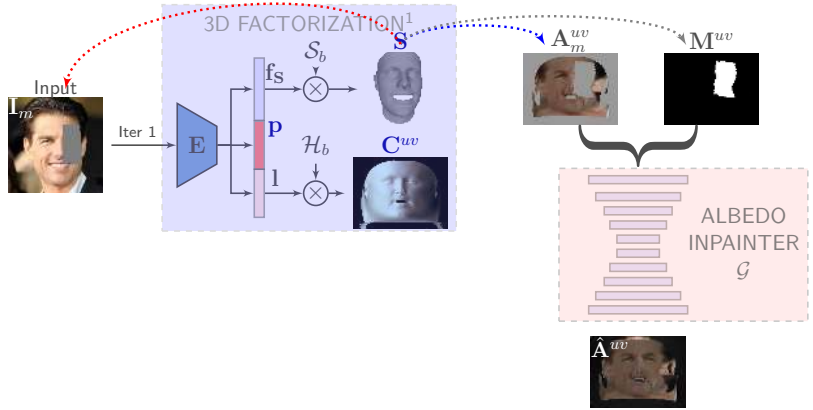
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



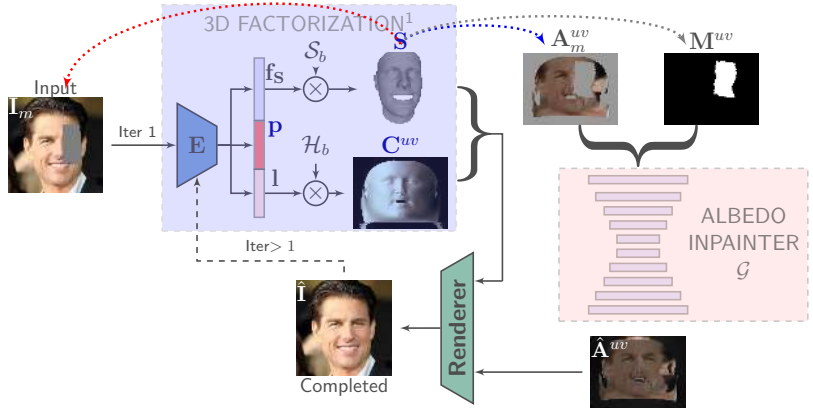
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



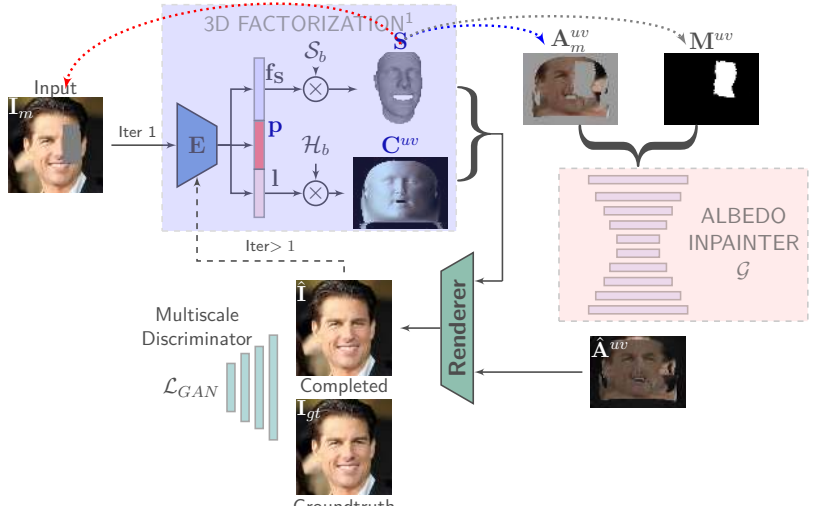
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



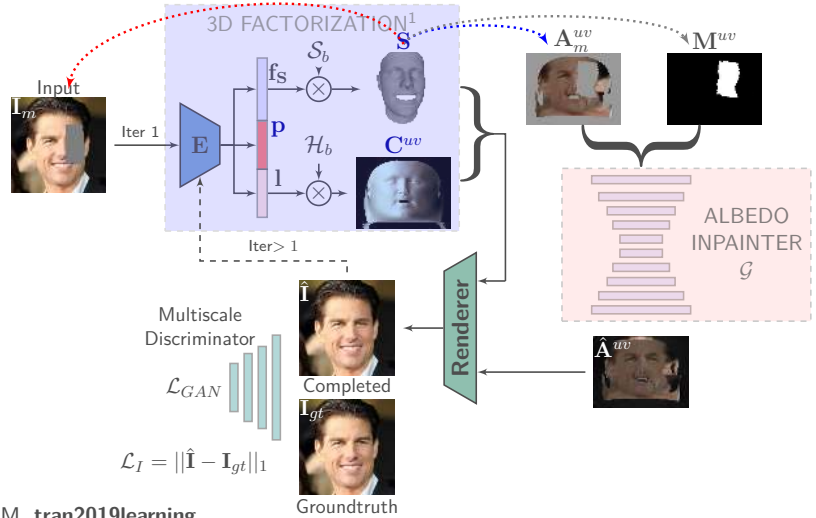
[1] Nonlinear 3DMM, **tran2019learning**

Architecture



[1] Nonlinear 3DMM, **tran2019learning**

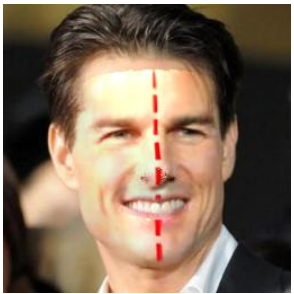
Architecture



[1] Nonlinear 3DMM, **tran2019learning**

Leveraging symmetry

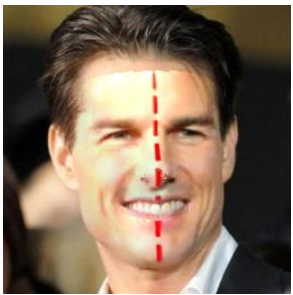
2D



Leveraging symmetry

2D

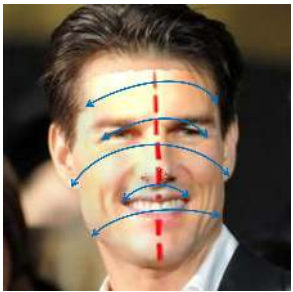
Variations in Illumination, Shape and Expression



Leveraging symmetry

2D

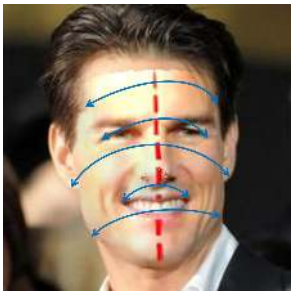
Variations in Illumination, Shape and Expression



Leveraging symmetry

Variations in Illumination, Shape and Expression

2D



vs.

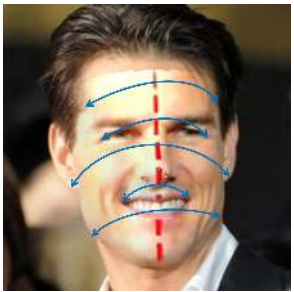
Albedo

Leveraging symmetry

2D

Albedo

Variations in Illumination, Shape and Expression



vs.

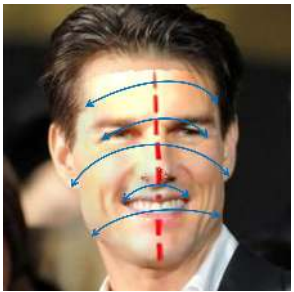
Disentangled from Illumination, Shape and Expression

Leveraging symmetry

2D

Albedo

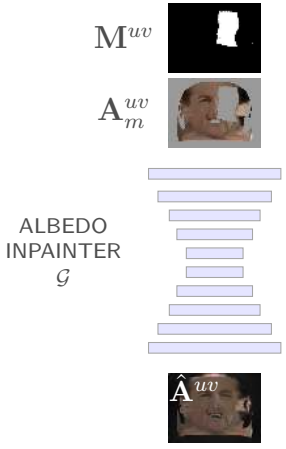
Variations in Illumination, Shape and Expression



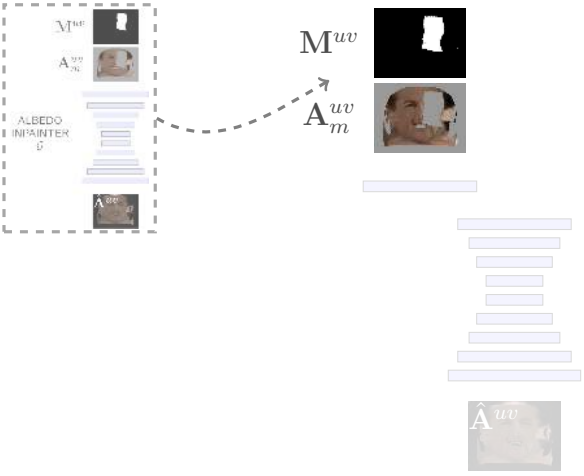
vs.

Disentangled from Illumination, Shape and Expression

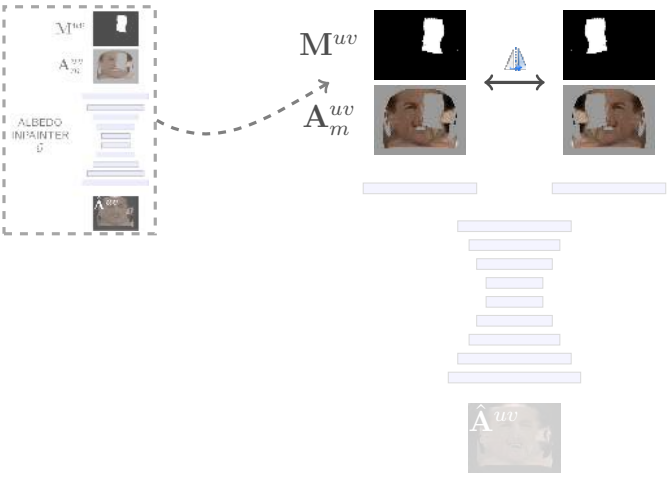
Sym-UNet: Albedo completion module that leverages symmetry



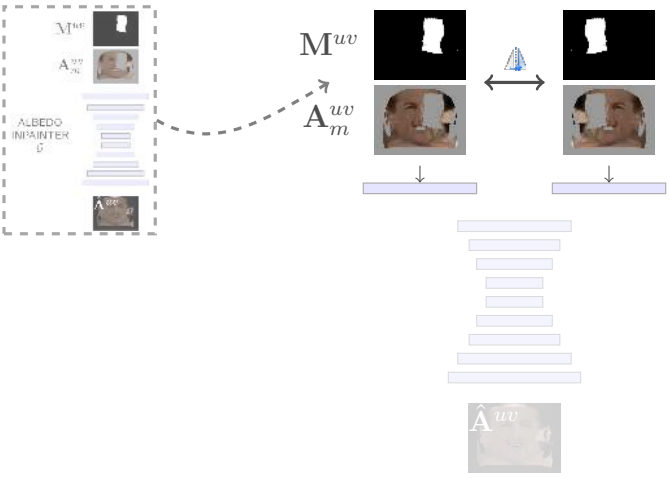
Sym-UNet: Albedo completion module that leverages symmetry



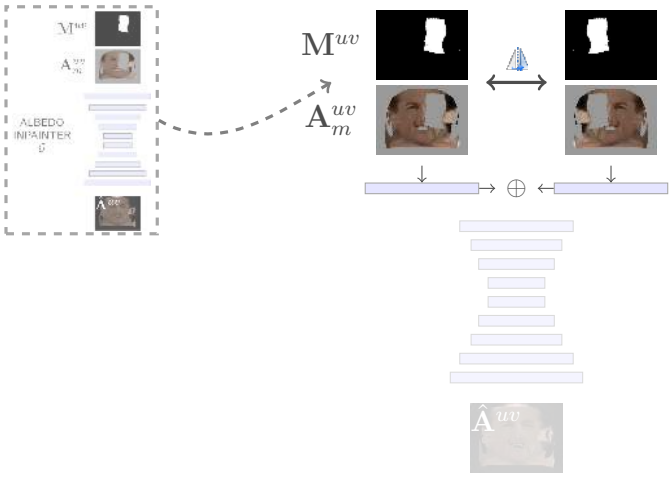
Sym-UNet: Albedo completion module that leverages symmetry



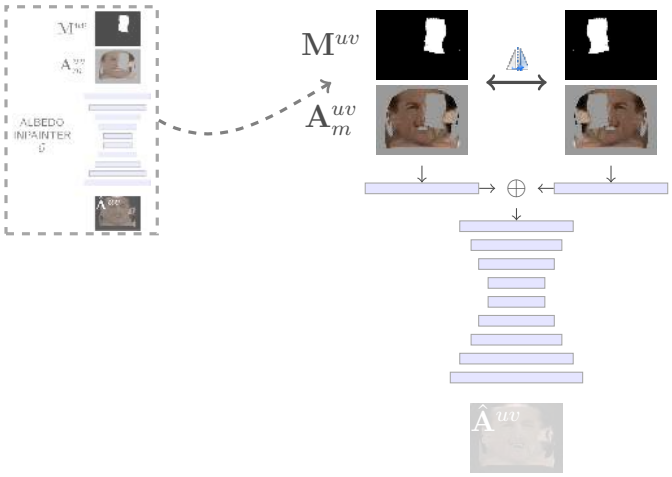
Sym-UNet: Albedo completion module that leverages symmetry



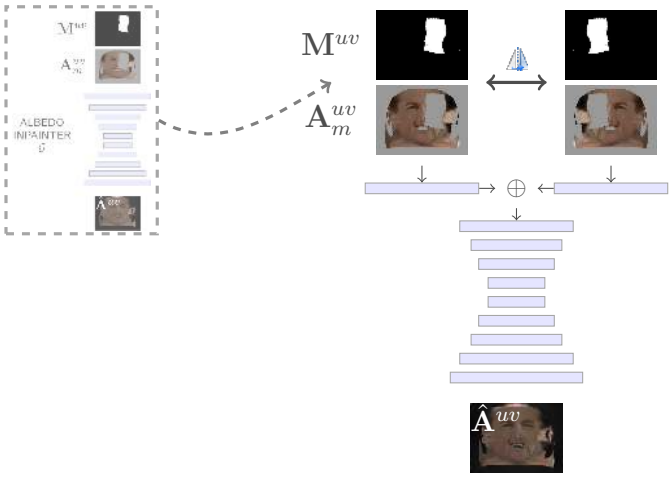
Sym-UNet: Albedo completion module that leverages symmetry



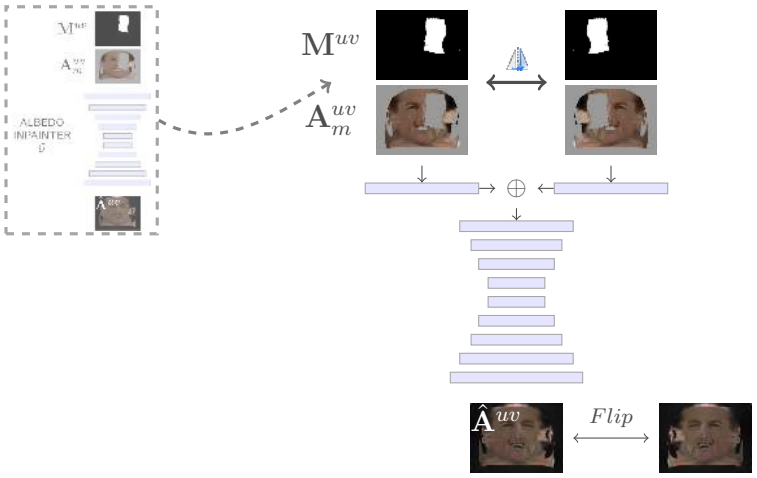
Sym-UNet: Albedo completion module that leverages symmetry



Sym-UNet: Albedo completion module that leverages symmetry

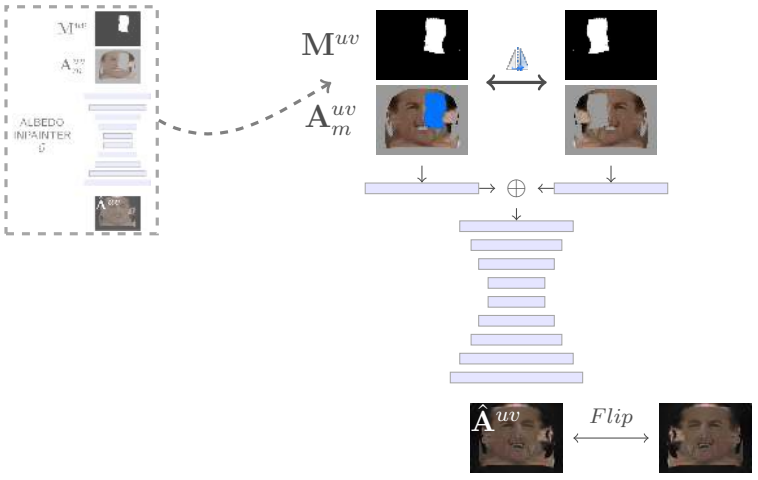


Sym-UNet: Albedo completion module that leverages symmetry



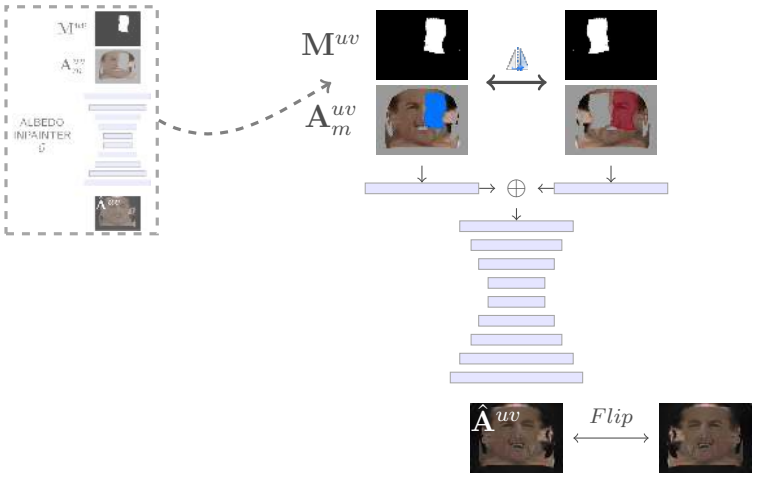
$$\mathcal{L}_{sym} = \mathbf{M}^{uv} \cdot \neg \mathbf{M}^{uv}_{flip} \cdot ||\hat{\mathbf{A}}^{uv} - \hat{\mathbf{A}}^{uv}_{flip}||_1$$

Sym-UNet: Albedo completion module that leverages symmetry



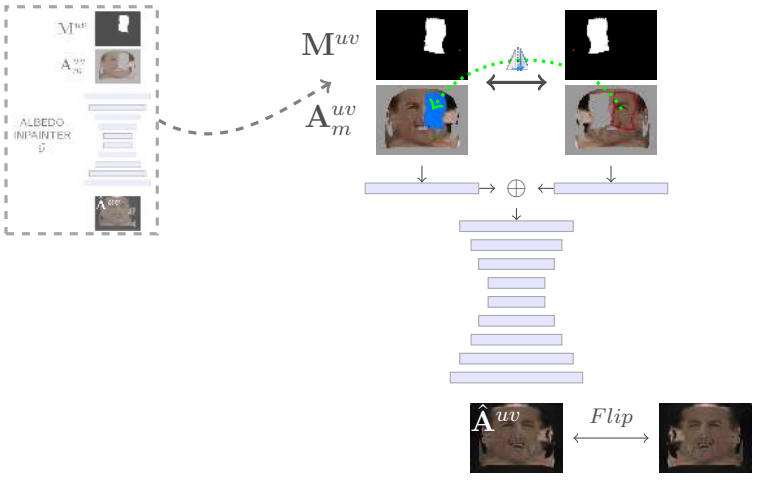
$$\mathcal{L}_{sym} = \mathbf{M}^{uv} \cdot \neg \mathbf{M}^{uv}_{flip} \cdot ||\hat{\mathbf{A}}^{uv} - \hat{\mathbf{A}}^{uv}_{flip}||_1$$

Sym-UNet: Albedo completion module that leverages symmetry



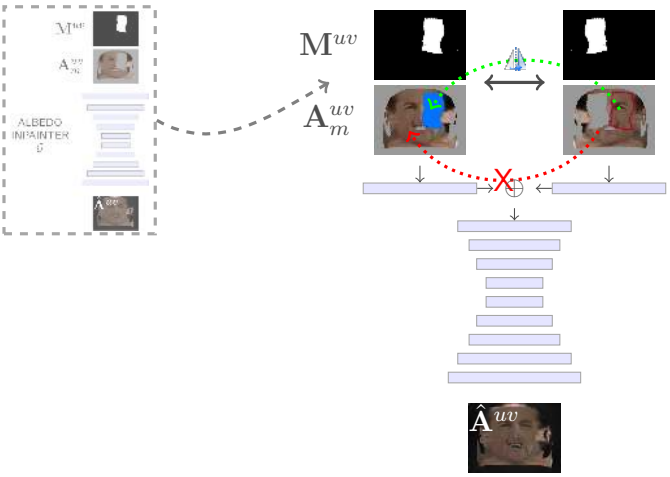
$$\mathcal{L}_{sym} = \mathbf{M}^{uv} \cdot \neg\mathbf{M}^{uv}_{flip} \cdot ||\hat{\mathbf{A}}^{uv} - \hat{\mathbf{A}}^{uv}_{flip}||_1$$

Sym-UNet: Albedo completion module that leverages symmetry

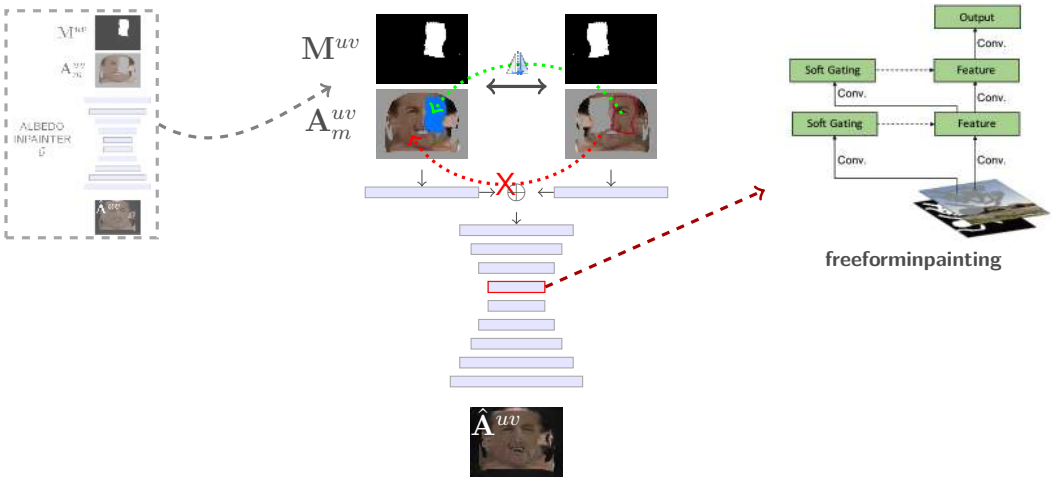


$$\mathcal{L}_{sym} = \mathbf{M}^{uv} \cdot \neg\mathbf{M}^{uv}_{flip} \cdot ||\hat{\mathbf{A}}^{uv} - \hat{\mathbf{A}}^{uv}_{flip}||_1$$

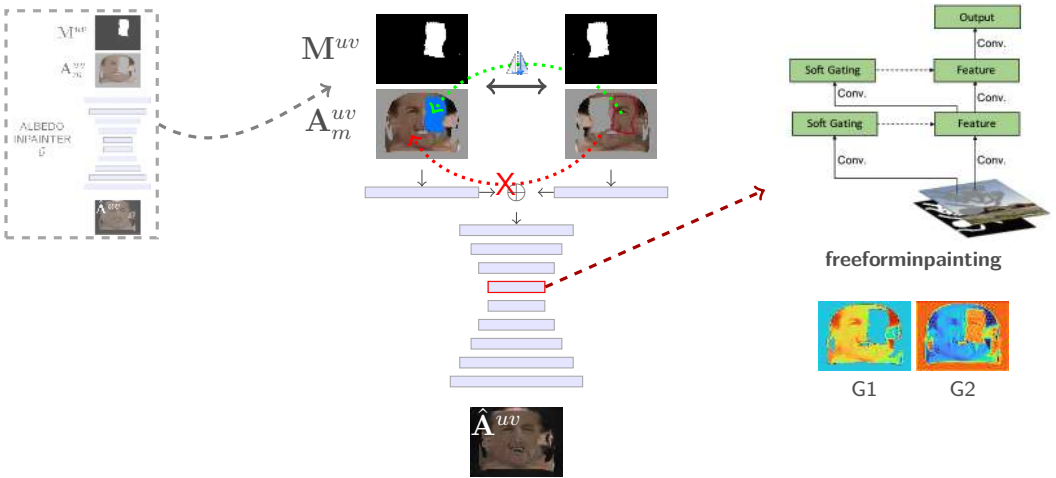
Sym-UNet: Albedo completion module that leverages symmetry



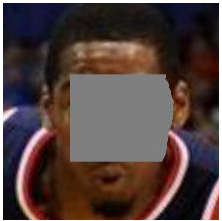
Sym-UNet: Albedo completion module that leverages symmetry



Sym-UNet: Albedo completion module that leverages symmetry



Qualitative evaluation - Dark complexion



Input

Qualitative evaluation - Dark complexion

Input

DeepFillv2

PIC

DSA

Qualitative evaluation - Dark complexion

Input

DeepFillv2

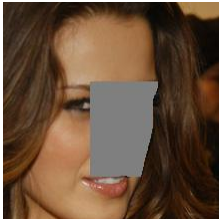
PIC

DSA

3DFaceFill (Ours)

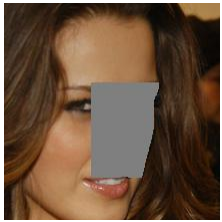
Groundtruth

Qualitative evaluation - Challenging poses

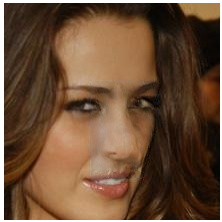


Input

Qualitative evaluation - Challenging poses



Input

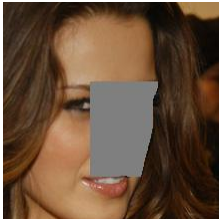


DeepFillv2

PIC

DSA

Qualitative evaluation - Challenging poses

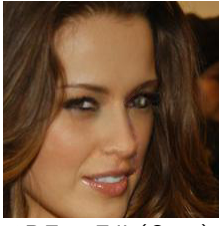


Input

DeepFillv2

PIC

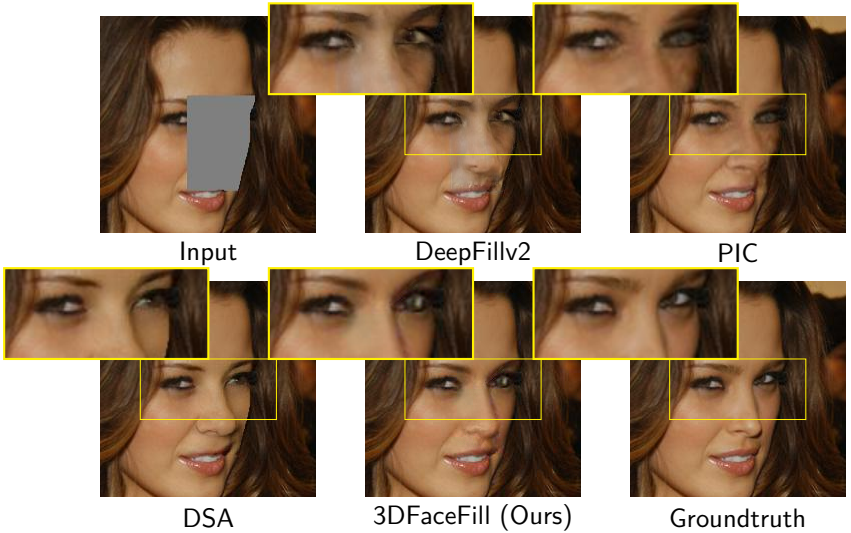
DSA



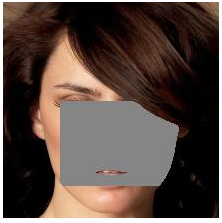
3DFaceFill (Ours)

Groundtruth

Qualitative evaluation - Challenging poses

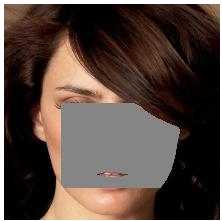


Qualitative evaluation - Shape deformation



Input

Qualitative evaluation - Shape deformation



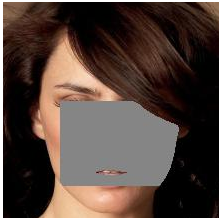
Input

DeepFillv2

PIC

DSA

Qualitative evaluation - Shape deformation



Input

DeepFillv2

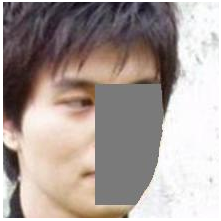
PIC

DSA

3DFaceFill (Ours)

Groundtruth

Qualitative evaluation - Eye-gaze symmetry



Input

Qualitative evaluation - Eye-gaze symmetry

Input

DeepFillv2

PIC

PConv

Qualitative evaluation - Eye-gaze symmetry

Input

DeepFillv2

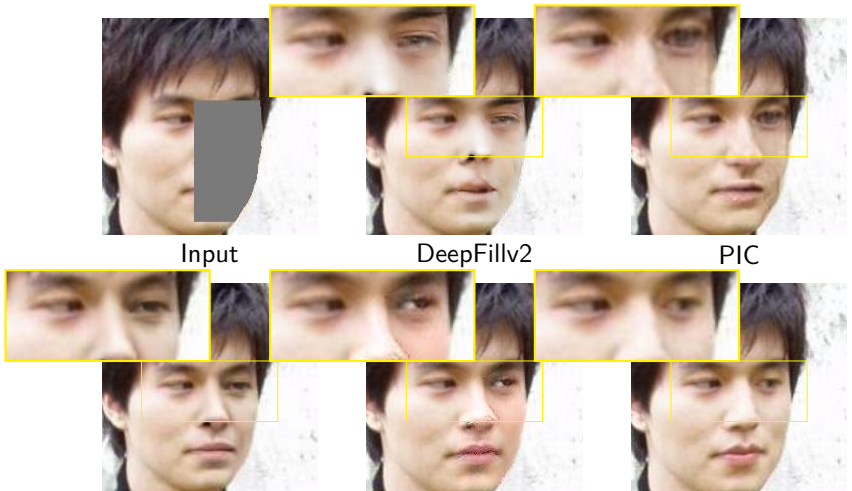
PIC

PConv

3DFaceFill (Ours)

Groundtruth

Qualitative evaluation - Eye-gaze symmetry



Qualitative evaluation - Kids

Input

Qualitative evaluation - Kids

Input

DeepFillv2

PIC

PConv

Qualitative evaluation - Kids

Input

DeepFillv2

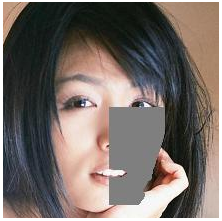
PIC

PConv

3DFaceFill (Ours)

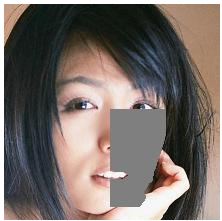
Groundtruth

Qualitative evaluation - Illumination variations



Input

Qualitative evaluation - Illumination variations



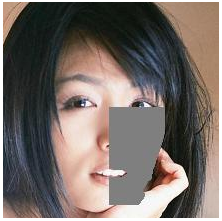
Input

DeepFillv2

PIC

SymmFCNet

Qualitative evaluation - Illumination variations



Input

DeepFillv2

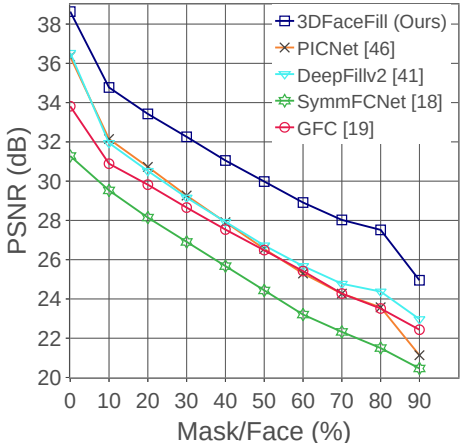
PIC

SymmFCNet

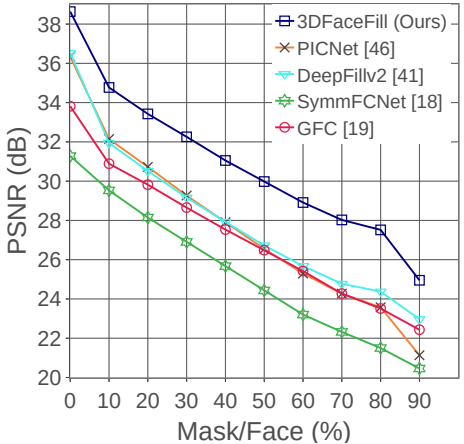
3DFaceFill (Ours)

Groundtruth

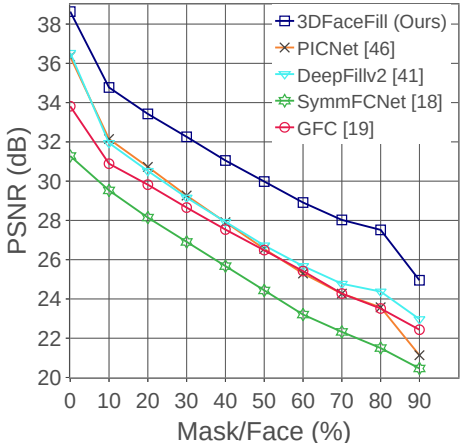
Quantitative evaluation on Celeb-A dataset



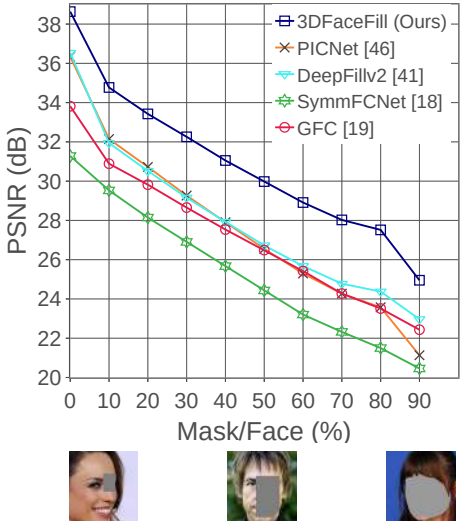
Quantitative evaluation on Celeb-A dataset



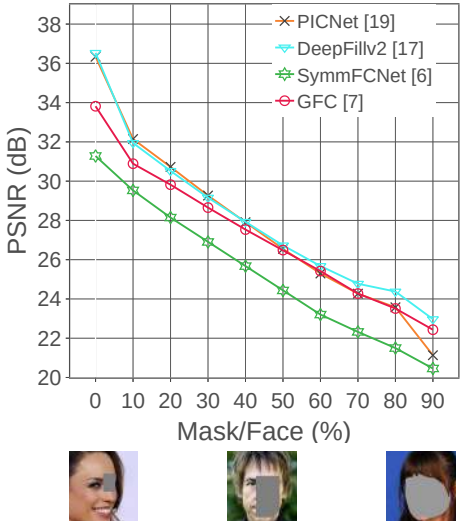
Quantitative evaluation on Celeb-A dataset



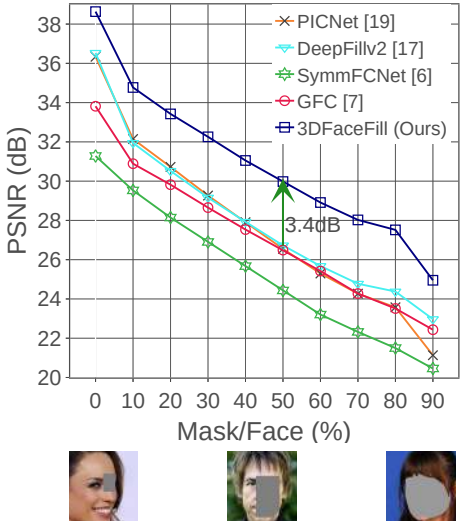
Quantitative evaluation on Celeb-A dataset



Quantitative evaluation on Celeb-A dataset

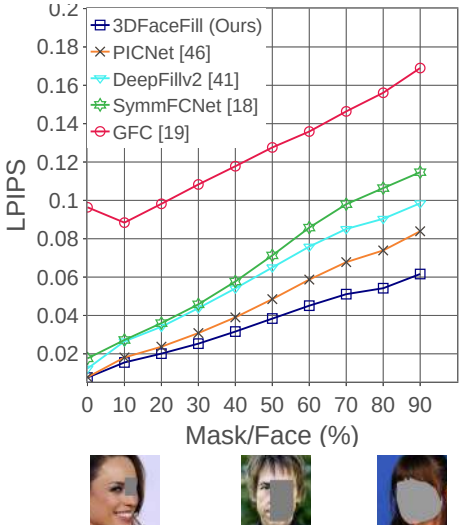


Quantitative evaluation on Celeb-A dataset



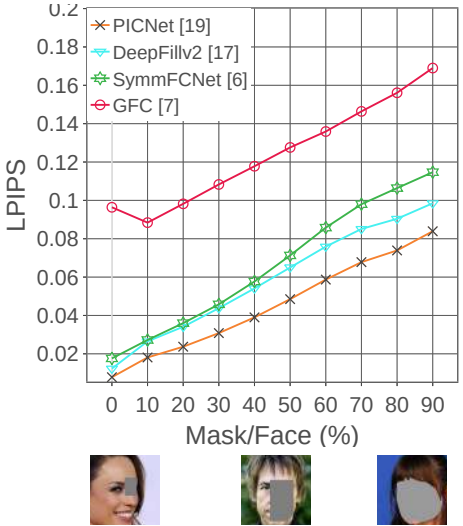
Quantitative evaluation Celeb-A dataset

LPIPS: Learned Perceptual Image Patch Similarity
lpips



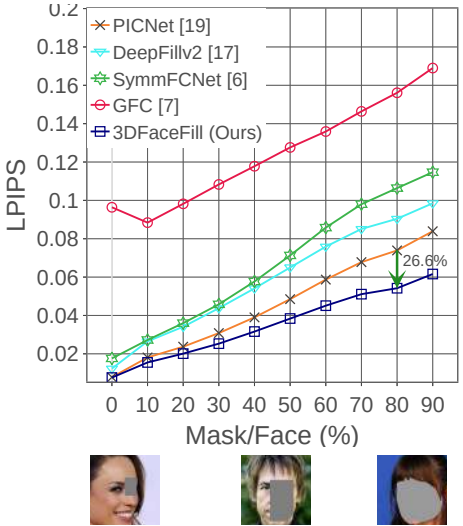
Quantitative evaluation Celeb-A dataset

LPIPS: Learned Perceptual Image Patch Similarity
lpips

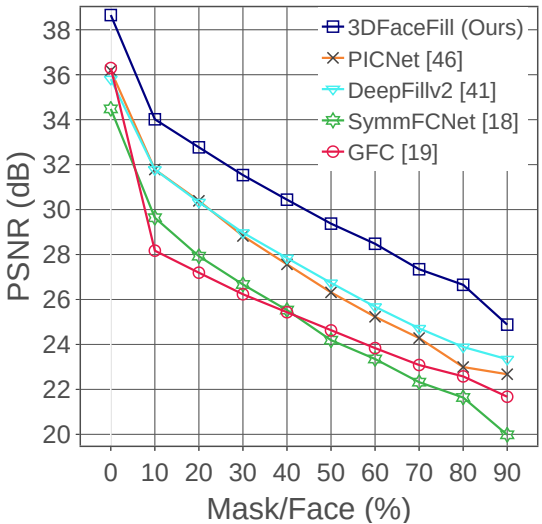


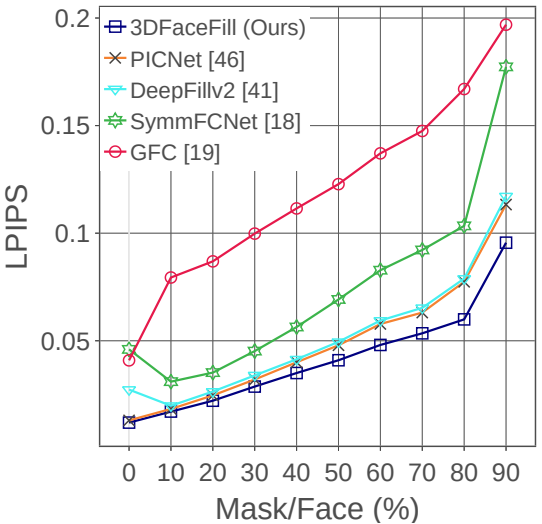
Quantitative evaluation Celeb-A dataset

LPIPS: Learned Perceptual Image Patch Similarity
lpips



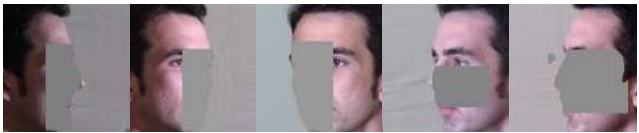
Quantitative evaluation - CelebA-HQ Dataset



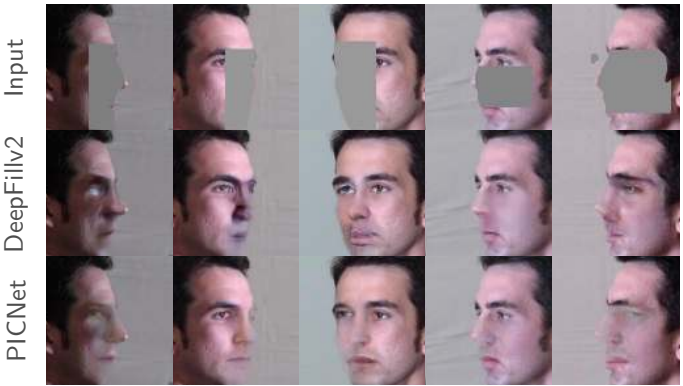


Qualitative evaluation - MultiPIE Pose

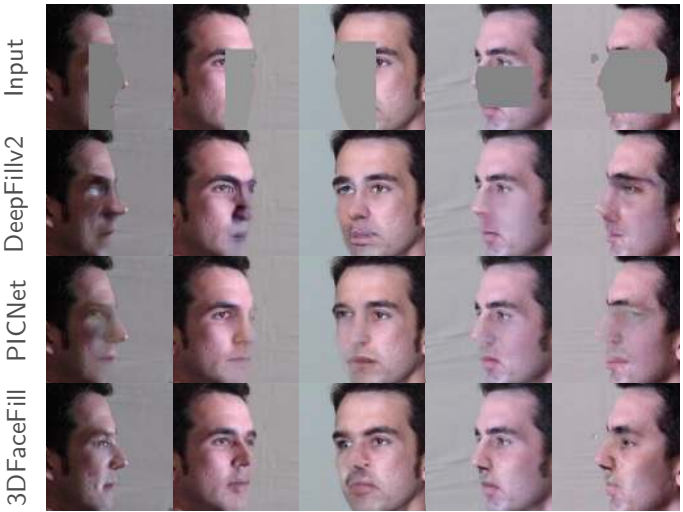
Input



Qualitative evaluation - MultiPIE Pose



Qualitative evaluation - MultiPIE Pose

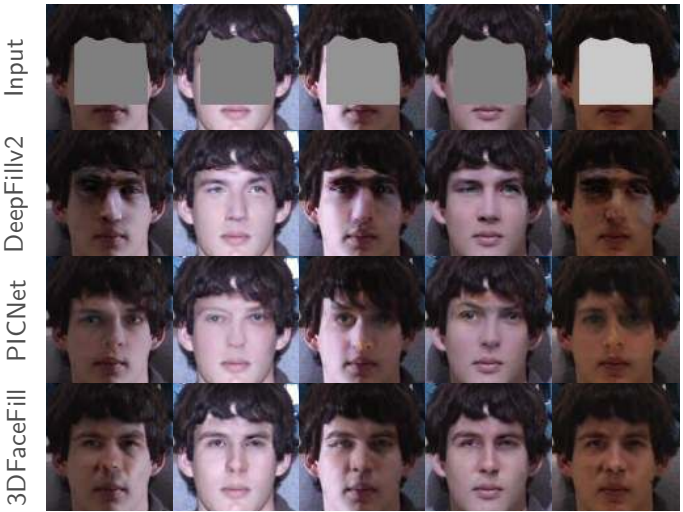


Qualitative evaluation - MultiPIE Illumination

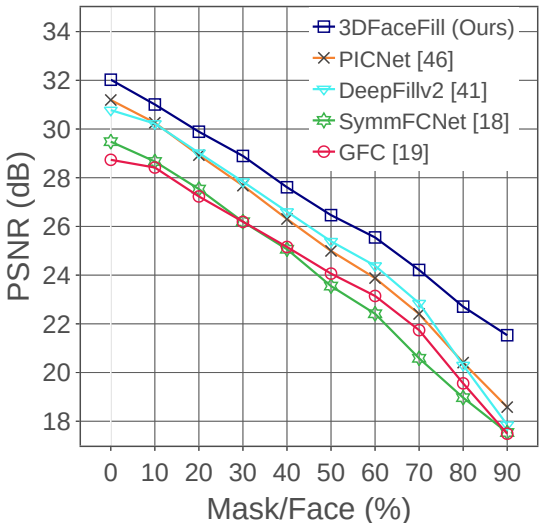
Input

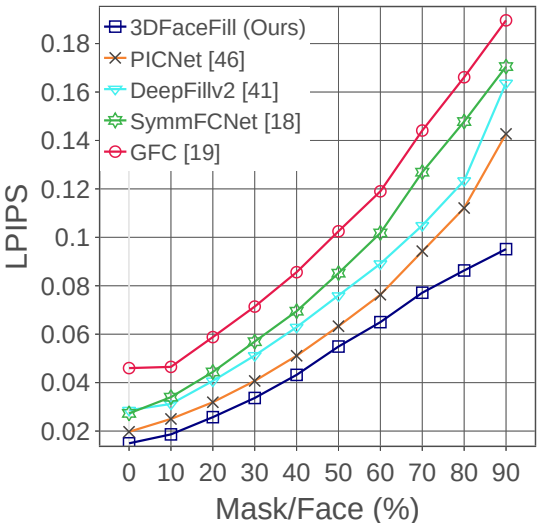
Qualitative evaluation - MultiPIE Illumination

Qualitative evaluation - MultiPIE Illumination

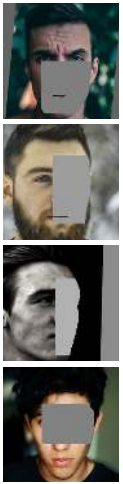


Quantitative evaluation - MultiPIE Pose and Illumination





Generalization performance on Internet faces

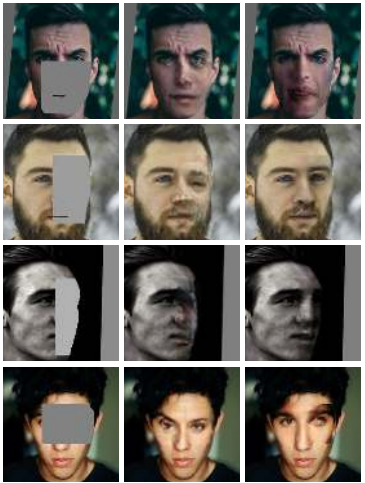


Input

Generalization performance on Internet faces

Input GFC

Generalization performance on Internet faces

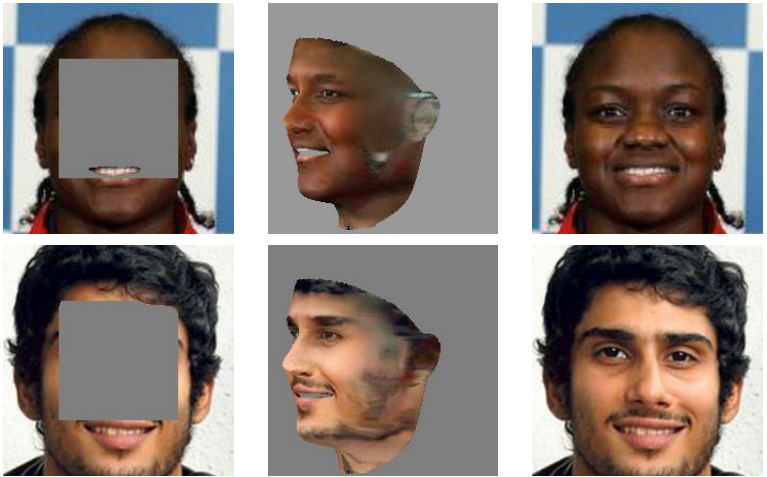


Generalization performance on Internet faces

Generalization performance on Internet faces

Generalization performance on Internet faces

3D completion and view synthesis



Input | Completed views | Ground Truth

Effect of iterative refinement

Input

Effect of iterative refinement

Input

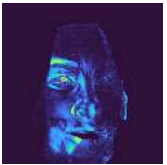
Groundtruth

Effect of iterative refinement


Input


Groundtruth

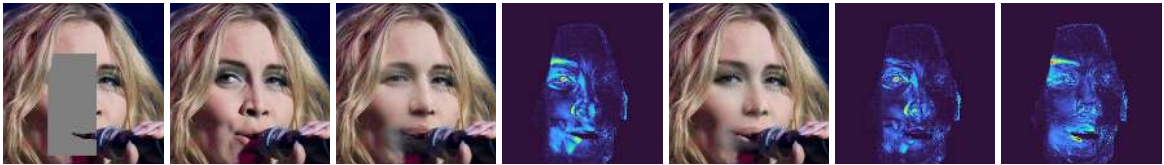

Iter1


Iter1-GT

Effect of iterative refinement

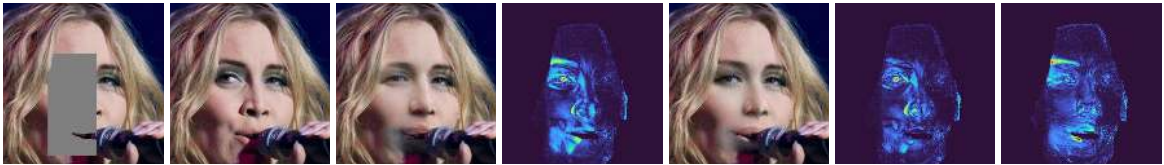
Input Groundtruth Iter1 Iter1-GT Iter2 Iter2-GT

Effect of iterative refinement

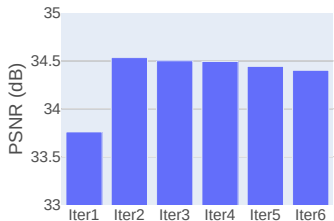


Input | Groundtruth | Iter1 | Iter1-GT | Iter2 | Iter2-GT | Iter2-Iter1

Effect of iterative refinement

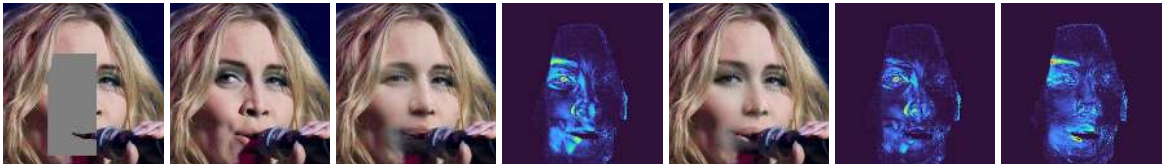


Input | Groundtruth | Iter1 | Iter1-GT | Iter2 | Iter2-GT | Iter2-Iter1

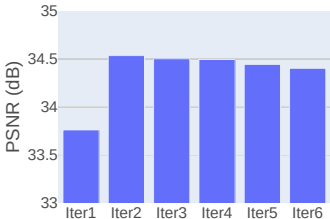


- Coarse inpainting leads to finer 3D modelling, which leads to finer inpainting

Effect of iterative refinement

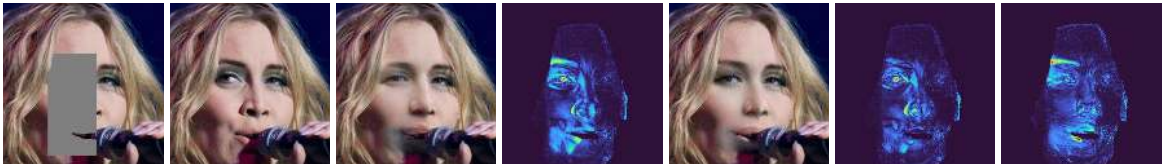


Input | Groundtruth | Iter1 | Iter1-GT | Iter2 | Iter2-GT | Iter2-Iter1

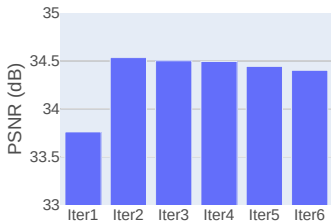


- Coarse inpainting leads to finer 3D modelling, which leads to finer inpainting
- We hypothesize that the stagnation (or slight dip) after Iter2 is because of worse pose estimation

Effect of iterative refinement



Input | Groundtruth | Iter1 | Iter1-GT | Iter2 | Iter2-GT | Iter2-Iter1



- Coarse inpainting leads to finer 3D modelling, which leads to finer inpainting
- We hypothesize that the stagnation (or slight dip) after Iter2 is because of worse pose estimation
- Two iterations are sufficient

Effect of Sym-UNet and the symmetry loss

Input

Groundtruth

Effect of Sym-UNet and the symmetry loss

Input

Groundtruth

NoSym Model

Effect of Sym-UNet and the symmetry loss

Input

Groundtruth

NoSym Model

NoSym+Attn

Effect of Sym-UNet and the symmetry loss

Input

Groundtruth

NoSym Model

NoSym+Attn

Full Model

Full-NoSym

Effect of Sym-UNet and the symmetry loss

Input

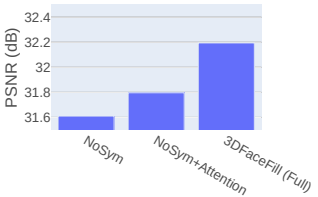
Groundtruth

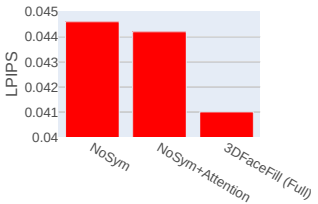
NoSym Model

NoSym+Attn

Full Model

Full-NoSym





Effect of Sym-UNet and the symmetry loss

Input

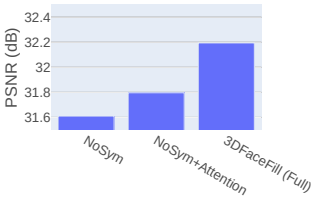
Groundtruth

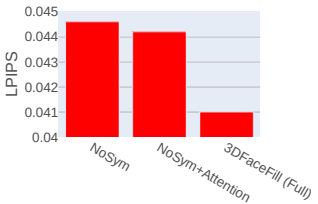
NoSym Model

NoSym+Attn

Full Model

Full-NoSym





- Attention helps in image completion

Effect of Sym-UNet and the symmetry loss

Input

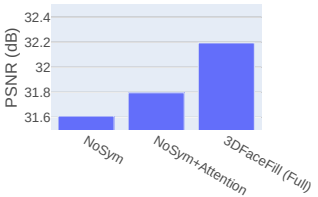
Groundtruth

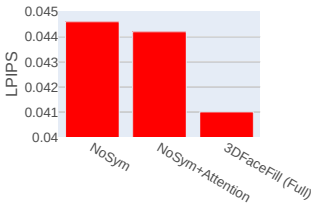
NoSym Model

NoSym+Attn

Full Model

Full-NoSym





- Attention helps in image completion
- For FC, symmetry is a stronger prior than attention

Conclusions

- Face completion using explicit 3D priors leads to geometrically and photometrically better results

Conclusions

- Face completion using explicit 3D priors leads to geometrically and photometrically better results
- Our method efficiently leverages facial symmetry

Conclusions

- Face completion using explicit 3D priors leads to geometrically and photometrically better results
- Our method efficiently leverages facial symmetry
- Qualitative and quantitative improvement in face completion under diverse conditions of shape, pose, illumination, *etc*

Conclusions

- Face completion using explicit 3D priors leads to geometrically and photometrically better results
- Our method efficiently leverages facial symmetry
- Qualitative and quantitative improvement in face completion under diverse conditions of shape, pose, illumination, *etc*
- Limitations:
  - Completes only the face region, excluding inner mouth

Conclusions

- Face completion using explicit 3D priors leads to geometrically and photometrically better results
- Our method efficiently leverages facial symmetry
- Qualitative and quantitative improvement in face completion under diverse conditions of shape, pose, illumination, *etc*
- Limitations:
  - Completes only the face region, excluding inner mouth
  - Resolution is limited by the resolution of the 3D mesh

Explicit 3D and Symmetry Priors

Thank you.