Supplemental Materials: PATMAT

A. Dataset

We used images of seven public figures and celebrities to train our models. For evaluation, we used 35 images per person, with the number of training images shown in table 3. Dwayne Johnson and Oprah Winfrey were the only people which had images with glasses and sunglasses. For that reason, we manually split them into train and test sets to have images with glasses in both sets. For the rest of the celebrities, we randomly split them into train and test sets.

Celebrity	Training size	Test size
Barack Obama	47	35
Dwayne Johnson	46	35
Oprah Winfrey	43	35
Scarlett Johansson	45	35
Taylor Swift	41	35
Joe Biden	42	35
Michelle Obama	40	35

Table 3: Training and test sizes used in our experiments.

B. Qualitative Study

We created three surveys as follows. survey **A** had 19 participants and 35 questions. Each question gave participants the choice to pick an image that best represented one of the seven people we used in our study (table 3). Both given options were inpaintings of the same image, with the same mask, one using PATMAT-S and one using PATMAT-C. We asked 5 questions per identity, totalling 35 questions.

These questions relied on the participants knowing and remembering what each figure looked like. We provided 3 real images of each person, not appearing anywhere else in our study and survey questions, as a reference. 62% of the answers picked PATMAT-C as a better representation of the celebrities.

Survey **B** had 28 participants and 10 questions. Each question had 4 different images of a person. Each participant had to pick what they thought was an inpainted image from three real images and one inpainted image using PATMAT-C. Survey **C** followed the same structure as survey **B**, with 21 participants, except in each question, the participants were given two images of a person; one real and one inpainted. In Survey **B**, 43% of the answers were able to tell an image was inpainted. In survey **C**, 56% were able to correctly pick the inpainted image.

Figure 9: Additional error analysis examples of PATMAT.

Figure 10: Tuning MAT without anchors for Oprah, inpaints images with glasses like features.

C. Additional Error Analysis

In figure 9, we show additional scenarios where PATMAT fails to properly inpaint an image. Figure 9-**A** shows examples where PATMAT attempts to generate / reconstruct sunglasses, yet fails to do so in a convincing fashion. Our experiments showed that this happens when the reference images do not contain the same type of glasses, hence PATMAT inpaints the missing part of the glasses with what is has seen in the references. Figure 9-**B** shows examples of PATMAT failing to properly align the inpainted area with the local and non-local priors.

D. Tuning Without Anchors

Anchors give structure to tuning MAT. We showed how using one anchor point to inpaint images with and without glasses will result in glasses-like features spreading to other images. Training with no anchors will have a similar effect, where in each iteration, random s_u noise-style codes are picked for tuning with reference image x_i. Without any structure to s_u codes, separating the features, figure 10 shows similar behaviour to what we saw in figure 2.

E. Additional Qualitative Results

Figure 11 shows additional inpainting examples using PATMAT-S, PATMAT-C and other competing methods.

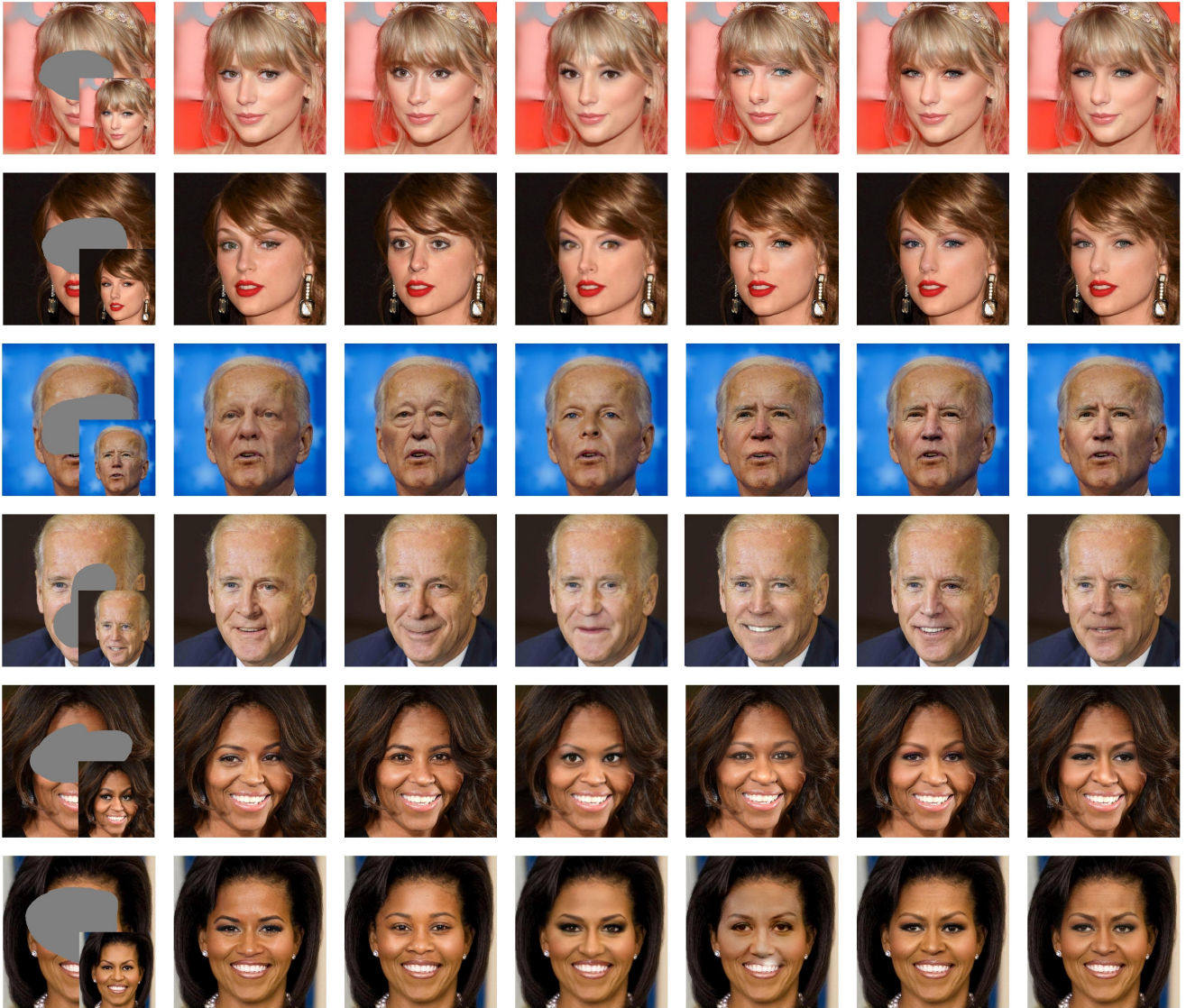| Masked Input	MAT	CoModGAN	RePaint	MyStyle	PATMAT-S	PATMAT-C

Figure 11: Additional qualitative results. Please zoom in for details.

Figure 11: Additional qualitative results. Please zoom in for details.