Supplementary Material for Image Super-Resolution with Guarantees via Conformalized Generative Models

Anonymous Author(s)
Affiliation
Address
email

A Proofs

Proof of Theorem 2.1. This proof is done via the standard conformal risk control argument [Angelopoulos et al., 2022].

Consider the "lifted" threshold

$$t_\alpha^{(n+1)} = \sup \left\{ t \in \mathbb{R} \cup \{+\infty\} : \frac{1}{n+1} \sum_{i=1}^{n+1} \sup_{p;[\sigma(X_i)]_p \le t} D_p\left(\mu(X_i), Y_i\right) \le \alpha \right\},$$

which, as opposed to their "unlifted" counterpart t_α, leverage the $(n+1)$-th sample as well and does not include the $3/(n+1)$ term.

Note that it is certainly the case that $t_\alpha^{(n+1)} \ge t_\alpha$. Moreover, note that the fidelity error is monotone with t, and so it suffices to show that the fidelity function with $t^{(n+1)}$ is upper bounded by α.

Let Z_* be the multiset of the samples $(X_i, Y_i)_{i=1}^n$ – i.e., a random variable representing the samples, but with their order discarded. Hence, upon conditioning on Z_*, all the randomness that remains is that of the order of the samples. It then follows:

$$\mathbb{E}\left[\sup_{p;[\sigma(X_{n+1})]_p \le t_\alpha^{(n+1)}} D_p(\mu(X_{n+1}), Y_{n+1}) \middle| Z_* \right] = \frac{1}{n+1} \sum_{i=1}^{n+1} \sup_{p;[\sigma(X_{n+1})]_p \le t_\alpha^{(n+1)}} D_p(\mu(X_i), Y_i),$$

and by the definition of $t_\alpha^{(n+1)}$, this is upper bounded by α. Thus

$$\begin{aligned} \mathbb{E}\left[\sup_{p;[\sigma(X_{n+1})]_p \le t_\alpha^{(n+1)}} D_p(\mu(X_{n+1}), Y_{n+1}) \right] &= \mathbb{E}_{Z_*}\left[\mathbb{E}\left[\sup_{p;[\sigma(X_{n+1})]_p \le t_\alpha^{(n+1)}} D_p(\mu(X_{n+1}), Y_{n+1}) | Z_* \right]\right] \\ &\le \mathbb{E}_{Z_*}[\alpha] = \alpha, \end{aligned}$$

which concludes the proof. □

Proof of Proposition 3.1. The PSNR we are bounding is given by

$$\begin{aligned} \mathbb{E}[\mathrm{PSNR}\left(\mu(X_{n+1}), Y_{n+1} | M_\alpha(X_{n+1})\right)] &:= \mathbb{E}\left[10 \log_{10} \frac{(\max_{p \in M_\alpha(X_{n+1})} [Y_{n+1}]_p)^2}{|M_\alpha(X_{n+1})|^{-1} \sum_{p \in M_\alpha(X_{n+1})} \left([\mu(X_{n+1})]_p - Y_p\right)^2} \right] \\ &= 20 \mathbb{E}\left[\log_{10} \frac{\max_{p \in M_\alpha(X_{n+1})} [Y_{n+1}]_p}{\sqrt{|M_\alpha(X_{n+1})|^{-1} \sum_{p \in M_\alpha(X_{n+1})} \left([\mu(X_{n+1})]_p - Y_p\right)^2}} \right]. \end{aligned}$$

Now, by Jensen's Inequality and standard properties of logarithms,

$$\begin{aligned}
&20\mathbb{E}\left[\log_{10}\frac{\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p}{\sqrt{|M_\alpha(X_{n+1})|^{-1}\sum_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)^2}}\right]\\
&=20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p-\log_{10}\sqrt{|M_\alpha(X_{n+1})|^{-1}\sum_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)^2}\right]\right)\\
&=20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p\right]-\mathbb{E}\left[\log_{10}\sqrt{|M_\alpha(X_{n+1})|^{-1}\sum_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)^2}\right]\right)\\
&\geq 20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p\right]-\log_{10}\mathbb{E}\left[\sqrt{|M_\alpha(X_{n+1})|^{-1}\sum_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)^2}\right]\right);
\end{aligned}$$

And, because the RMSE is upper bounded by the maximum error, we get that

$$\begin{aligned}
&20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p\right]-\log_{10}\mathbb{E}\left[\sqrt{|M_\alpha(X_{n+1})|^{-1}\sum_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)^2}\right]\right)\\
&\geq 20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p\right]-\log_{10}\mathbb{E}\left[\sup_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)\right]\right);
\end{aligned}$$

And, by Theorem 2.1,

$$\begin{aligned}
&20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p\right]-\log_{10}\mathbb{E}\left[\sup_{p\in M_\alpha(X_{n+1})}\left([\mu(X_{n+1})]_p-Y_p\right)\right]\right)\\
&\geq 20\left(\mathbb{E}\left[\log_{10}\max_{p\in M_\alpha(X_{n+1})}[Y_{n+1}]_p\right]-\log_{10}\alpha\right)\geq -20\log_{10}\alpha,
\end{aligned}$$

where the last step holds as long as all pixel values are in $[0,1]$.

$\square$

Proof of 3.2. We effectively want to bound the supremum of the expected fidelity error as the leaked data is allowed to alter freely. For convenience, let $\sup_{\text{leaked}}$ denote the supremum over all possible values of the leaked samples $(X_i,Y_i)_{i=n_{\text{new}}+1}^{n}$ (and $\inf_{\text{leaked}}$ the corresponding infimum).

Note that the error function is decreasing on the selected parameter t and continuous. Hence:

$$\sup_{\text{leaked}}\mathbb{E}\left[\sup_{p\in M_\alpha(X)}D_p(Y,\widehat{Y})\right]\leq\mathbb{E}\left[\sup_{\text{leaked}}\sup_{p\in M_\alpha(X)}D_p(Y,\widehat{Y})\right]=\mathbb{E}\left[\sup_{p;[\sigma(X)]_p\leq\sup_{\text{leaked}}t_\alpha}D_p(Y,\widehat{Y})\right],$$

and in turn

$$
\begin{aligned}
\sup_{\text{leaked}} t_\alpha &= \sup_{\text{leaked}} \sup\left\{t \in \mathbb{R}\cup\{+\infty\} : \frac{1}{n+1}\sum_{i=1}^{n} \sup_{p;[\sigma(X_i)]_p \le t} D_p(Y_i, \mu(X_i)) + \frac{3}{n+1} \le \alpha\right\} \\
&\le \sup\left\{t \in \mathbb{R}\cup\{+\infty\} : \inf_{\text{leaked}} \frac{1}{n+1}\sum_{i=1}^{n} \sup_{p;[\sigma(X_i)]_p \le t} D_p(Y_i, \mu(X_i)) + \frac{3}{n+1} \le \alpha\right\} \\
&= \sup\Bigg\{t \in \mathbb{R}\cup\{+\infty\} : \frac{1}{n+1}\sum_{i=1}^{n_{\text{new}}} \sup_{p;[\sigma(X_i)]_p \le t} D_p(Y_i, \mu(X_i)) \\
&\qquad + \inf_{\text{leaked}} \frac{1}{n+1}\sum_{i=n_{\text{new}}+1}^{n} \sup_{p;[\sigma(X_i)]_p \le t} D_p(Y_i, \mu(X_i)) + \frac{3}{n+1} \le \alpha\Bigg\} \\
&= \sup\left\{t \in \mathbb{R}\cup\{+\infty\} : \frac{1}{n+1}\sum_{i=1}^{n_{\text{new}}} \sup_{p;[\sigma(X_i)]_p \le t} D_p(Y_i, \mu(X_i)) + \frac{3}{n+1} \le \alpha\right\} \\
&= \sup\left\{t \in \mathbb{R}\cup\{+\infty\} : \frac{1}{n_{\text{new}}+1}\sum_{i=1}^{n_{\text{new}}} \sup_{p;[\sigma(X_i)]_p \le t} D_p(Y_i, \mu(X_i)) + \frac{3}{n_{\text{new}}+1} \le \alpha \cdot \frac{n+1}{n_{\text{new}}+1}\right\}.
\end{aligned}
$$

Note that this corresponds to doing our calibration procedure only on the new data but with altered fidelity level $\alpha \cdot (n+1)/(n_{\text{new}}+1) = \alpha \cdot (n_{\text{new}} + n_{\text{leaked}} + 1)/(n_{\text{new}}+1)$, and so, by the same arguments as in Theorem 2.1,

$$
\mathbb{E}_{(X_i,Y_i)_{i=1}^{n+1}}\left[\sup_{p\in M_\alpha(X_{n+1})} D_p\left(\mu(X_{n+1}), Y_{n+1}\right)\right] \le \alpha \cdot \frac{n_{\text{new}} + n_{\text{leaked}} + 1}{n_{\text{new}}+1}.
$$

$\square$

B On [Kutiel et al., 2023]

The method of [Kutiel et al., 2023] produces, given a low-resolution image X in $[0,1]^{w\times h\times 3}$ and an initial estimate of uncertainty σ in $\mathbb{R}_{\geq 0}^{kw\times kh}$, a predicted high-resolution image $\hat{Y}$ in $[0,1]^{kw\times kh\times 3}$ along with a continuous 'confidence mask' M in $[0,1]^{kw\times kh}$, where higher values denote higher confidence at that region of the image. Ideally, this continuous mask would be such that it would satisfy an RCPS-like guarantee

$$\mathbb{P}\left[\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}[M]_{i,j}\cdot\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]\leq\alpha\right]\geq 1-\delta, \tag{1}$$

for some chosen α and δ. This goal is enunciated in Definition 2 in [Kutiel et al., 2023].[1] Their calibration procedure is then presented in Section 4.2 of their paper; it does the following:

1. For each $i=1,\ldots,n$, compute

$$M^{(\lambda)}:=\min\left\{\frac{\lambda}{1-[\sigma]_{i,j}+\epsilon},1\right\}$$
$$\lambda_i:=\max\left\{\lambda:\frac{1}{kw\,kh}\sum_{i,j}[M^{(\lambda)}]_{i,j}\cdot\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\leq\alpha\right\}$$

2. Compute $\lambda:=\mathrm{Quantile}_{1-\delta}(\lambda_1,\ldots,\lambda_n)$.
3. The resulting "calibrated" masks are produced by $M^{(\lambda)}$.

However, they do not prove that this procedure satisfies Equation 1, other than a passing mention at the end of their Section 4.2. Indeed, the guarantee actually does not hold: intuitively this should be fairly immediate:

1. They write[2] "Finally, λ is taken to be the $1-\delta$ quantile of $\{\lambda_k\}_{k=1}^{n}$, i.e. the maximal value for which at least δ fraction of the calibration set satisfies condition (5). Thus, assuming the calibration and test sets are i.i.d samples from the same distribution, the calibrated mask is guaranteed to satisfy Definition 2." However, their condition (5) states that

$$\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}[M]_{i,j}\cdot\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]\leq\alpha,$$

   with an expectation – this expectation is fundamentally absent in the calibration procedure.
2. Additionally, the quantile taken is a simple empirical quantile. But the guarantee we want is that the populational risk is bounded. To satisfy this, one would need to slightly tweak the quantile; this is analogous to how in RCPSs you would apply a concentration inequality (e.g. Hoeffding).

Indeed, we provide here a counterexample, in which their procedure does not satisfy their stated guarantee.

[1] The one difference is the presence of the normalization $\frac{1}{kw\,kh}$, which is absent in their paper (they refer to a simple 1-norm). However, we found it to be necessary in order for their method to function, and conjecture that it was accidentally omitted in their work.

[2] Modulo ajustments to notation

Example B.1. Suppose that σ is a deterministic 0 mask for all inputs. Then, to show that Equation 1 does not hold, we just need that

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}[M]_{i,j}\cdot\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]\leq\alpha\right]<1-\delta\\
\iff&\mathbb{P}\left[\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}[M]_{i,j}\cdot\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]>\alpha\right]>\delta\\
\iff&\mathbb{P}\left[[M]_{i,j}\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]>\alpha\right]>\delta\\
\iff&\mathbb{P}\left[\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]>\frac{\alpha}{[M]_{i,j}}\right]>\delta,
\end{aligned}$$

for some α and δ. Further suppose that there is some nonzero probability τ that $\frac{1}{kw\,kh}\sum_{i,j}\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|=0$ for all samples in the calibration set. Then, with this nonzero probability, $\lambda=1$ regardless of δ, implying that $[M]_{i,j}=1/(1+\epsilon)$. By chosing $\delta=\tau/2$ and $\alpha\leq\frac{1+\epsilon}{2}\mathbb{E}\left[\frac{1}{kw\,kh}\sum_{i,j}\left|[\hat{Y}]_{i,j}-[Y]_{i,j}\right|\right]$, we conclude.

C Pseudocodes

Concrete implementations of these algorithms can be found in our code.

Algorithm 1 Conformal mask calibration (computation of t_α) with dynamic programming

```
T ← all unique values of σ(X_1), ..., σ(X_n)
T ← Sort(T)
D ← values of D_p each entry of T                         ▷ can be computed jointly with the sorting
I ← indices of the original images for each entry of T    ▷ can be computed jointly with the sorting
R_i ← 0                                                   ▷ risk so far on each observation
R ← 3/(n+1)                                               ▷ total risk so far
t_⋆ ← −∞
l ← NA                                                    ▷ last threshold seen so far
for i = 1, ..., nd do
    if t_i ≠ l then
        if R ≤ α then
            t_⋆ ← t
        end if
    end if
    l ← t_i
    r ← max{R_{I_i}, D_i}
    R ← R − 1/(n+1) R_{I_i} + 1/(n+1) r
    R_i ← r
end for
return t_⋆
```

Algorithm 2 Conformal mask calibration (computation of t_α) with a brute force search

```
T ← all unique values of σ(X_1), ..., σ(X_n)
t_⋆ ← −∞
for t ∈ T do
    compute risk R ← 1/(n+1) Σ_{i=1}^{n} sup_{p;[σ(X_i)]_p ≤ t} D_p(Y_i, μ(X_i)) + 3/(n+1)
    if R ≤ α then
        t_⋆ ← max{t_⋆, t}
    end if
end for
return t_⋆
```

D Results on Image Colorization

Here we showcase a direct adaptation of our method for image colorization rather than super-resolution. The method works similarly, suggesting a broad applicability beyond our domain.

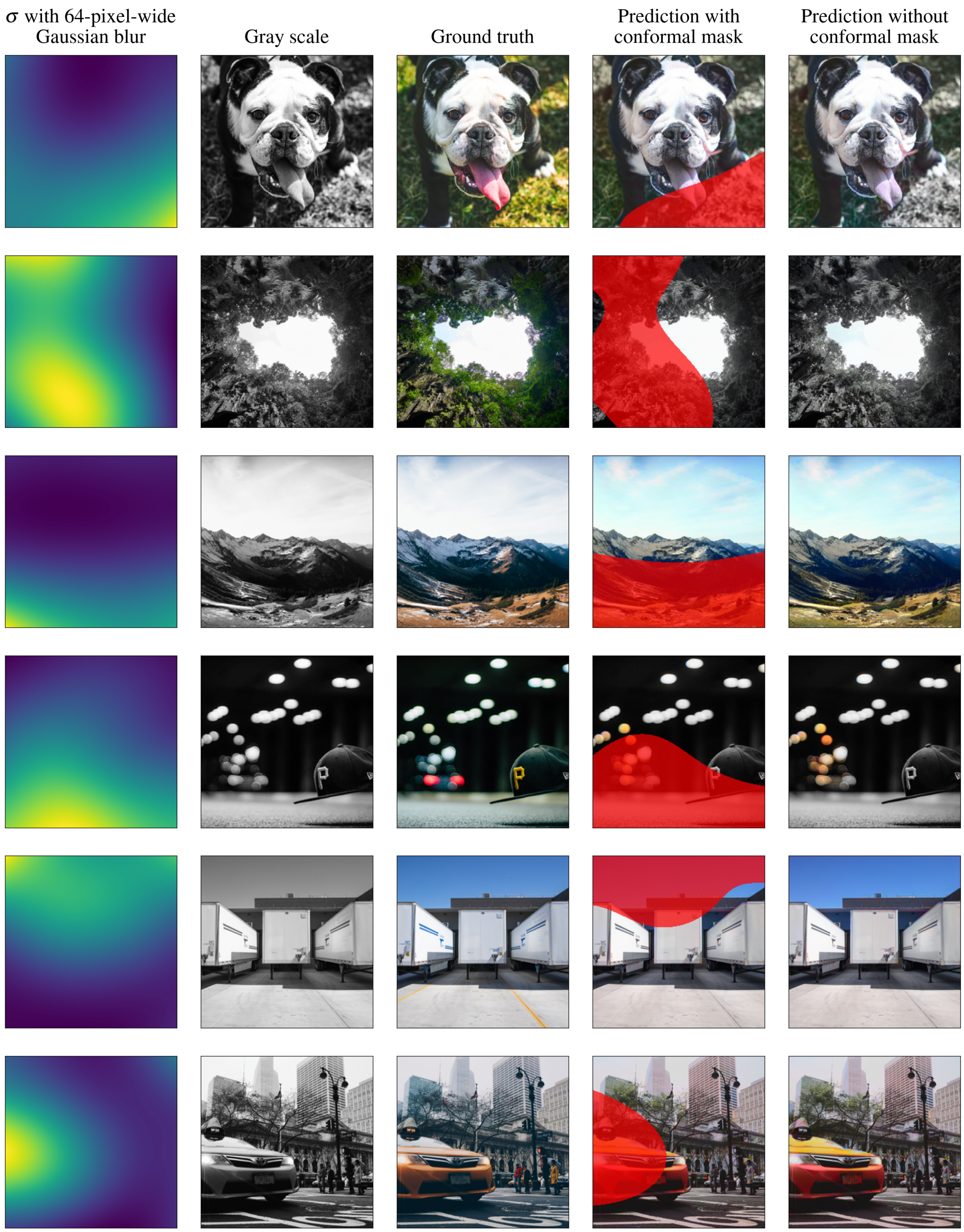

Figure 1: **Results of our method adapted for image colorization.** This figure presents a comparison of multiple color-restored images generated using the DDColor base model [Kang et al., 2023], alongside their corresponding conformal masks for $\alpha = 0.1$ and non-semantic D_p. Our conformal masks accurately highlight regions where color predictions significantly deviate from the ground truth, effectively capturing uncertainty related to hue variations, saturation inconsistencies, and lighting discrepancies.This demonstrates the versatility of our theoretical framework, indicating it can be directly extended to other image restoration tasks beyond super-resolution.

References

Anastasios Nikolas Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. ArXiv, abs/2208.02814, 2022. URL `https://api.semanticscholar.org/CorpusID:251320513`.

Xiaoyang Kang, Tao Yang, Wenqi Ouyang, Peiran Ren, Lingzhi Li, and Xuansong Xie. Ddcolor: Towards photo-realistic image colorization via dual decoders. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 328–338, 2023.

Gilad Kutiel, Regev Cohen, Michael Elad, Daniel Freedman, and Ehud Rivlin. Conformal prediction masks: Visualizing uncertainty in medical imaging. In TML4H, 2023. URL `https://api.semanticscholar.org/CorpusID:259311276`.