Conformal Prediction for Reliable Image Super-Resolution

Tapabrata Chakraborti TCHAKRABORTY@TURING.AC.UK
The Alan Turing Institute, London, United Kingdom.

Samiran Dey SMCSSD2661@IACS.RES.IN
Indian Association for the Cultivation of Science, Kolkata, India.

Editor: Khuong An Nguyen, Zhiyuan Luo, Harris Papadopoulos, Tuwe Löfström, Lars Carlsson and Henrik Boström

Abstract

Single image super-resolution (SISR) has been employed over a wide range of applications to enhance the visual quality and details of images. For training super-resolution (SR) models, low-resolution (LR) images are synthesized from the high-resolution (HR) images Dey et al. (2025). However, these artificial intelligence (AI) methods for SR (like diffusion based or generative adversarial models) have stochastic elements that are inherent to the learning process that can be mitigated but not avoided. Effectively this means that for the same input LR image, different instantiations of the generative process is expected to produce slightly different HR output images. While this might not be an issue for certain applications, and in fact might provide interesting variations in tasks like AI art generation, in certain other high stakes applications like medical image super-resolution, such variations need to be tightly controlled and rigorously quantified Chakraborti et al. (2025). After all, when superresolving a biomedical image (say radiology) one would ideally expect the output to be invariant for a patient if it is a static time-independent image. In fact, though the point of super-resolving a biomedical image is to provide the human expert (or indeed an equivalent AI system) the visual clarity to make a better evaluation, having degradation of clinical features or introduction of spurious morphological features would defeat the purpose, and potentially increase the chances of a false inference. Thus it is important in such high risk applications to predict uncertainty bounds for the generated images using a conformal prediction inspired estimate of maximum calibrated coverage.

Keywords: image super-resolution, biomedical radiology imaging, computed tomography (CT), diffusion models, generative adversarial networks (GAN), uncertainty quantification (UQ), generative artificial intelligence (AI), conformal prediction.

Conformal Bounds for Super-resolved Images

A function $u(x)$ is assumed whose values are higher for a high degree of uncertainty and lower when the uncertainty is less. We define $u(x) = \mid HR - G_{sr}(LR) \mid$ as the pixel-wise residual image between the generated SR, $G_{sr}(LR)$, and the ground truth HR image. The same model is next trained to learn $u(x)$ and generate the respective residual images for the model, referred as G_{res}. Next, we define a score function $s(HR, LR)$ that would act as an attention map to the residual image for predicting the bounds as $s(HR, LR) = \frac{u(LR)}{\hat{u}(LR)} = \frac{|HR - G_{sr}(LR)|}{G_{res}(LR)}$ where $\mathrm{G_{sr}}$ is the model trained to predict super-resolved images, $\mathrm{G_{res}}$ is the same model trained to generate its corresponding residual images and fraction implies pixel-wise division. Thus, the score function acts as a multiplicative correction factor or the attention map for the predicted residual image, that is $\hat{u}(LR) \cdot s(HR, LR) = \mid HR - G_{sr}(LR) \mid$. To approximate the value of the score function, we adopt the standard conformal prediction formulation. We assume an error rate of α is permissible. Thus essentially, the estimated value of score function ($\hat{a}$) is the $1 - \alpha$ quantile of the scores computed over a hold-out data of size N, which we call the calibration set. Introducing a small correction factor, Angelopoulos et al. (2025) defined $\hat{a} = \frac{\lceil (N+1)(1-\alpha) \rceil}{N}^{th}$ quantile of $s(HR, LR)$. Thus, we get the equation for the residual image as $u(x) = \hat{u}(LR) \cdot \hat{a} = G_{res}(LR) \cdot \hat{a}$. Hence, the predicted lower bound (I_{lb}) and upper bound (I_{ub}) images are respectively defined as $I_{lb} = G_{sr}(LR) - G_{res}(LR) \cdot \hat{a}$ and $I_{ub} = G_{sr}(LR) + G_{res}(LR) \cdot \hat{a}$ respectively. The estimated bounds provide a probable range of images that a model may generate. It is expected that this range has an image that is an exact match to the ground truth HR image, which we call coverage. Moreover, images within the predicted range will have different details enhanced at different levels, thus few images may be sampled and produced along with the SR image as alternates for clinicians to consider for diagnosis. To check for coverage of the HR images (HR_{cov}) offered by the predicted bounds, we check for the condition $\mathbb{P}\left(I_{lb} \leq HR \leq I_{ub}\right) \geq (1 - \alpha)$. Similarly, to check for coverage offered by the predicted bounds for different SR images (SR_{cov}) generated for the same HR using LR images having different levels of degradation, we check for the condition $\mathbb{P}\left(I_{lb} \leq G_{sr}(LR) \leq I_{ub}\right) \geq (1 - \alpha)$. SR coverage is also a measure of the bias of the model towards degradation in LR images corresponding to the same HR, and an estimate of the variations produced by the model in the anatomical structures and other minute details. For every image that satisfies the conditions, a unit point is awarded and the mean is obtained over the entire data to compute the coverage. The uncertainty (unc) is estimated as the mean of the L1 distance of the HR and the SR images corresponding to the same HR from the upper and lower bounds: $unc = \sum_{X_i} [(I_{lb} - X_i) + (I_{ub} - X_i)]/2$, where X consists of the HR image and the SR images corresponding to the same HR image generated from LR images varying only in the level of degradation. However, when super-resolving a real image, in the absence of a reference HR image the uncertainty is computed as the difference between the boundary images, and in such cases, coverage can not be calculated.

The contributions of this work can be listed as follows: 1) Residual image learning by the models to estimate a range of generated images that is most probable to contain an exact match to the ground truth image; 2) Provision to suggest alternate enhanced images within the predicted bounds, to the clinicians for better diagnosis of patients; 3) Estimation of uncertainty and coverage of generative medical image super-resolution models to contribute in clinical decision support system for personalised healthcare. To the best of our knowledge, this is the first work that predicts a probable range of images that a super-resolution model is capable of generating using estimated boundary images and hence, provides a measure of uncertainty. Note that this abstract presents the mathematical setup only, while further details and experimental results are included in Dey et al. (2025).

References

Anastasios N. Angelopoulos et al. Theoretical foundations of conformal prediction. *To be published by Cambridge University Press (preprint: arXiv:2411.11824)*, 2025.

Tapabrata Chakraborti et al. Personalised uncertainty quantification in artificial intelligence. *Nature Machine Intelligence*, 7:522–530, 2025.

Samiran Dey et al. Uncertainty estimation using boundary prediction for medical image super-resolution. *Computer Vision and Image Understanding*, 256(104349), 2025.