

| INT | Modality | Time (s) | PSNR(dB) / IoU(%) \uparrow |
|--------------|-----------|------------------|------------------------------|
| \times | Audio | 23.05 | 48.38 ± 3.50 |
| | Image | 345.22 | 36.09 ± 2.51 |
| | Megapixel | 16.78K | 31.82 |
| | 3D Shape | 144.58 | 97.07 ± 0.84 |
| \checkmark | Audio | 15.76 (-31.63%) | 48.15 ± 3.39 |
| | Image | 211.04 (-38.88%) | 36.97 ± 3.59 |
| | Megapixel | 11.87K (-29.26%) | 33.01 |
| | 3D Shape | 93.19 (-35.54%) | 96.68 ± 0.83 |