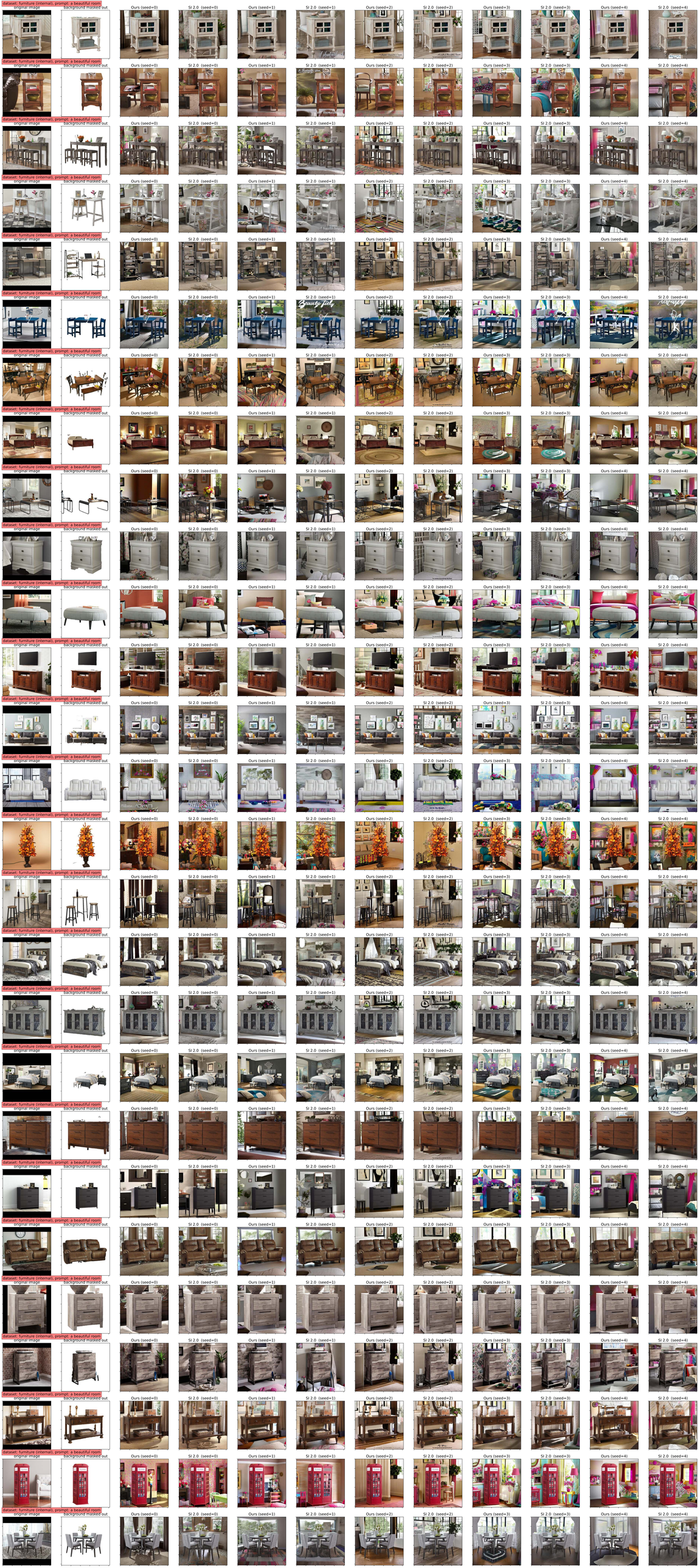
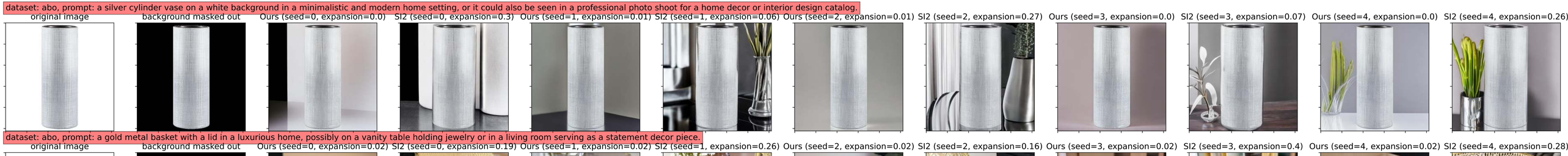

dataset: furniture (internal), prompt: a beautiful room
original image
background masked out
Ours (seed=0)
SI 2.0 (seed=0)
Ours (seed=1)
SI 2.0 (seed=1)
Ours (seed=2)
SI 2.0 (seed=2)
Ours (seed=3)
SI 2.0 (seed=3)
Ours (seed=4)
SI 2.0 (seed=4)

dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.27)
Ours (seed=1, expansion=0.02)
SI2 (seed=1, expansion=0.27)
Ours (seed=2, expansion=0.02)
SI2 (seed=2, expansion=0.08)
Ours (seed=3, expansion=0.02)
SI2 (seed=3, expansion=0.17)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.25)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.15)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.32)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.25)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.2)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.23)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.25)
Ours (seed=1, expansion=0.02)
SI2 (seed=1, expansion=0.26)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.24)
Ours (seed=3, expansion=0.09)
SI2 (seed=3, expansion=0.28)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.25)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.14)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.14)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.19)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.21)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.41)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.28)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.1)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.02)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.16)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.24)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.02)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.21)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.19)
Ours (seed=3, expansion=0.02)
SI2 (seed=3, expansion=0.15)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.15)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.3)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.05)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.1)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.39)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.14)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.16)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.12)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.09)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.21)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.1)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.14)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.08)
Ours (seed=3, expansion=0.1)
SI2 (seed=3, expansion=0.24)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.29)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.01)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.02)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.18)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.14)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.25)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.03)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.1)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.13)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.47)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.26)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.12)
Ours (seed=1, expansion=0.04)
SI2 (seed=1, expansion=0.27)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.13)
Ours (seed=3, expansion=0.04)
SI2 (seed=3, expansion=0.15)
Ours (seed=4, expansion=0.03)
SI2 (seed=4, expansion=0.27)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.13)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.12)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.1)
Ours (seed=3, expansion=0.03)
SI2 (seed=3, expansion=0.16)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.1)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.12)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.08)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.07)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.27)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.15)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.18)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.05)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.11)
Ours (seed=3, expansion=0.03)
SI2 (seed=3, expansion=0.11)
Ours (seed=4, expansion=0.04)
SI2 (seed=4, expansion=0.18)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.08)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.12)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.03)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.26)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.1)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.04)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.16)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.1)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.0)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.17)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.03)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.44)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.09)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.02)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.17)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.18)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.06)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.11)
Ours (seed=3, expansion=0.03)
SI2 (seed=3, expansion=0.06)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.14)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.01)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.04)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.1)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.11)
dataset: abo, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.06)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.12)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.02)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.15)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.09)

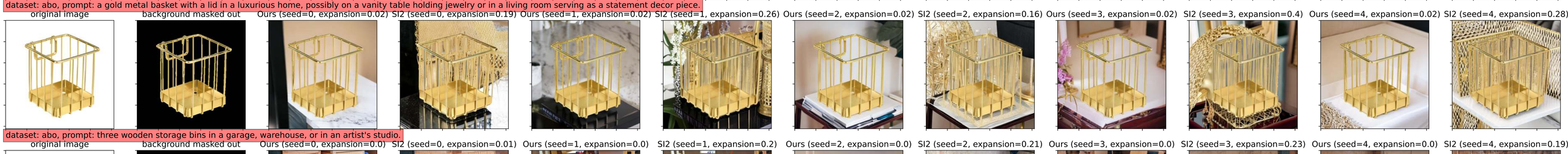
dataset: abo, prompt: a silver cylinder vase on a white background in a minimalistic and modern home setting, or it could also be seen in a professional photo shoot for a home decor or interior design catalog.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.3) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.06) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.27) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.07) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.26)



dataset: abo, prompt: a gold metal basket with a lid in a luxurious home, possibly on a vanity table holding jewelry or in a living room serving as a statement decor piece.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.19) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.4) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.28)

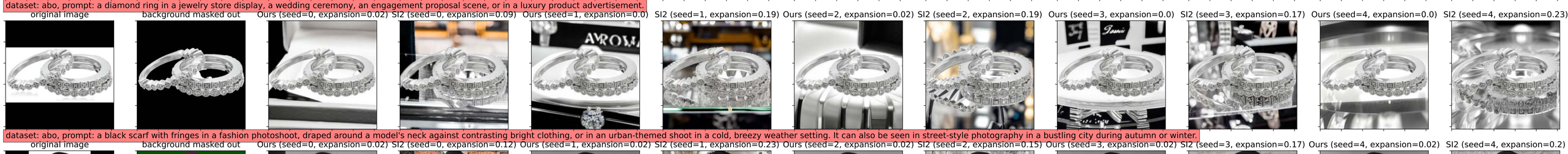


dataset: abo, prompt: three wooden storage bins in a garage, warehouse, or in an artist's studio.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.01) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.21) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.23) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)

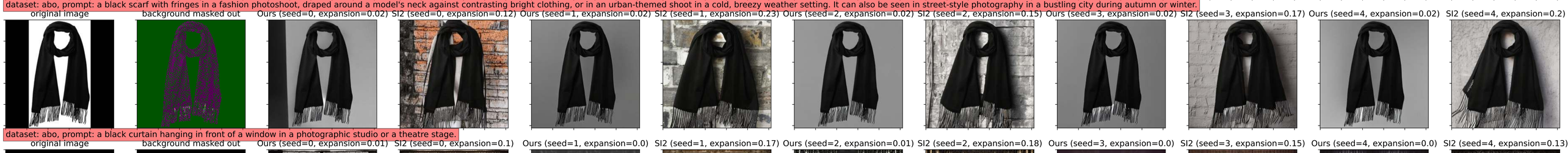
dataset: abo, prompt: a diamond ring in a jewelry store display, a wedding ceremony, an engagement proposal scene, or in a luxury product advertisement.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.19) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.23)



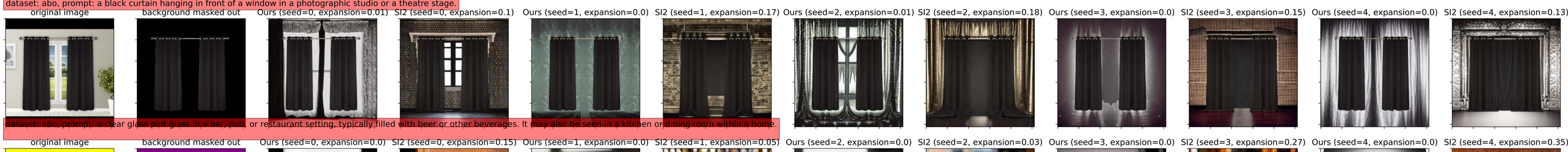
dataset: abo, prompt: a black scarf with fringes in a fashion photoshoot, draped around a model's neck against contrasting bright clothing, or in an urban-themed shoot in a cold, breezy weather setting. It can also be seen in street-style photography in a bustling city during autumn or winter.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.23) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.2)



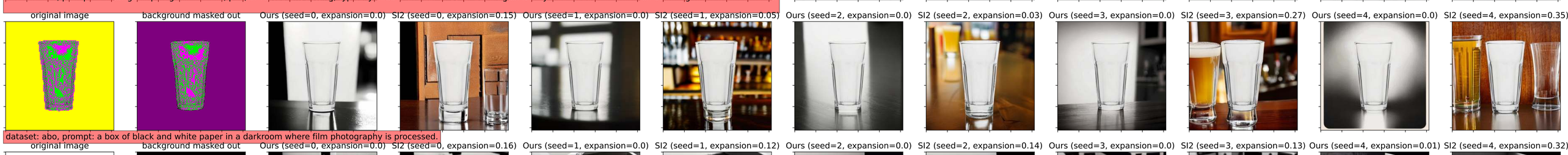
dataset: abo, prompt: a black curtain hanging in front of a window in a photographic studio or a theatre stage.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.1) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)



dataset: abo, prompt: a clear glass pint glass in a bar, pub, or restaurant setting, typically filled with beer or other beverages. It may also be seen in a kitchen or dining room within a home.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.05) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.03) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.27) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.35)



dataset: abo, prompt: a box of black and white paper in a darkroom where film photography is processed.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.32)

dataset: abo, prompt: women's black combat boots in a fashion photoshoot, city street style photography, or a military-themed photo session.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.06) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.15)

dataset: abo, prompt: a statement necklace in a fashion photo shoot or in a high-end jewelry store display.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.03) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.03) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.12)

dataset: abo, prompt: a gold ring with multi colored stones in a jewelry store display, a wedding ceremony, a high-end fashion photoshoot, or a luxurious vanity table.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.67) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.05) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.04) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

dataset: abo, prompt: three plastic containers with lids in a kitchen setting, being used for food storage.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.02) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

dataset: abo, prompt: a pair of men's green and blue sneakers in a scene such as a sports gym, a jogging track, a men's wardrobe, or a shoe store.

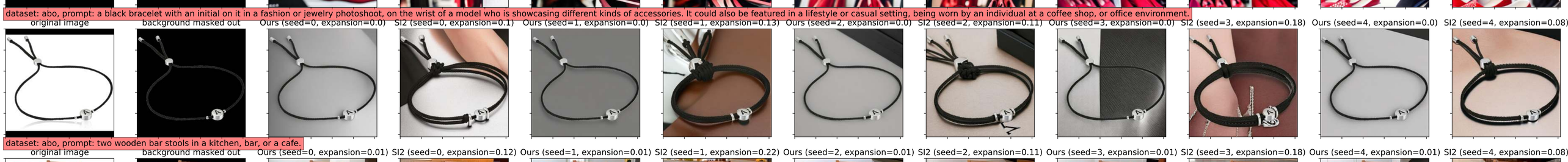
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.03) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.11) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.24) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.23)

dataset: abo, prompt: a pair of red flat shoes in a fashionable retail store, a shoe boutique, a closet, fashion show, or even in a casual street style themed photoshoot.

original image | background masked out | Ours (seed=0, expansion=0.04) | SI2 (seed=0, expansion=0.08) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.05) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.04) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.05) | SI2 (seed=4, expansion=0.16)

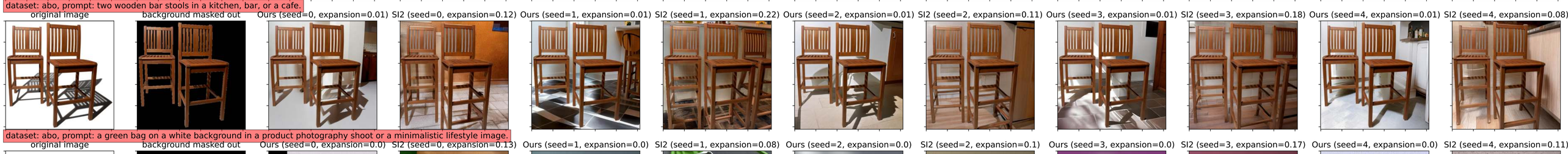
dataset: abo, prompt: a black bracelet with an initial on it in a fashion or jewelry photoshoot, on the wrist of a model who is showcasing different kinds of accessories. It could also be featured in a lifestyle or casual setting, being worn by an individual at a coffee shop, or office environment.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.1) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.18) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.08)



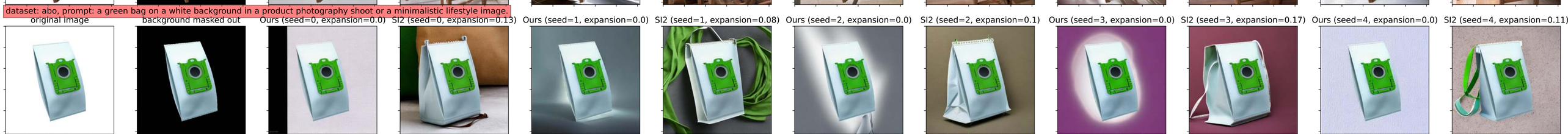
dataset: abo, prompt: two wooden bar stools in a kitchen, bar, or a cafe.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.18) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.08)



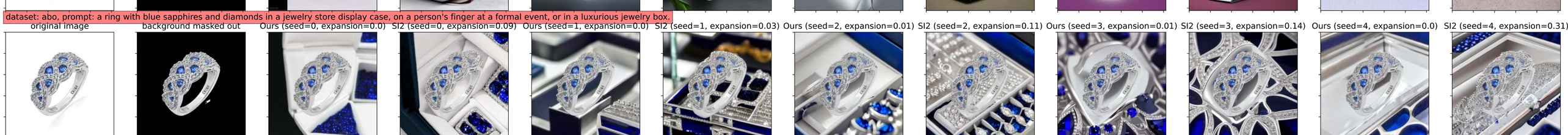
dataset: abo, prompt: a green bag on a white background in a product photography shoot or a minimalistic lifestyle image.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.08) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)



dataset: abo, prompt: a ring with blue sapphires and diamonds in a jewelry store display case, on a person's finger at a formal event, or in a luxurious jewelry box.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.03) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.31)



dataset: abo, prompt: a pair of black and yellow straps in a gym setting, particularly on the part of exercise or resistance training equipment.

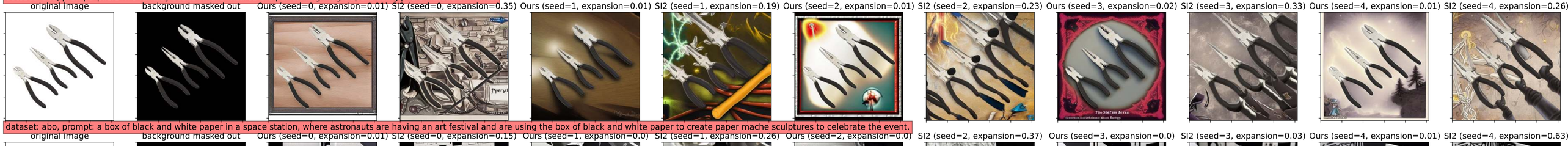
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.1) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.21) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.06) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.24)

dataset: abo, prompt: steak on a plate with arugula and croutons in a desert scene where a mirage of a fine dining restaurant appears amidst the endless dunes of sand.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.46) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.31) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.35) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.2)

dataset: abo, prompt: pliers in a fantasy scene where wizards are performing magic spells using pliers instead of wands.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.35) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.19) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.23) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.33) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.26)

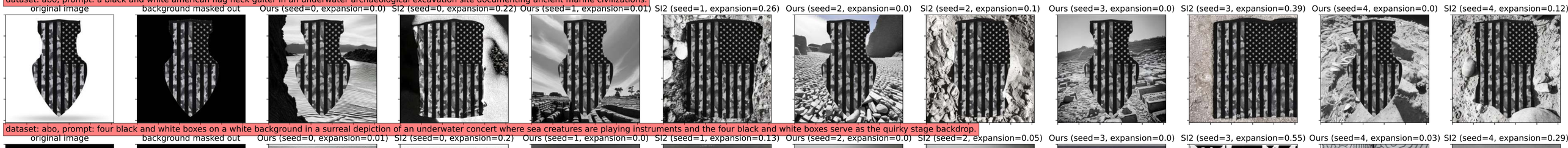


dataset: abo, prompt: a box of black and white paper in a space station, where astronauts are having an art festival and are using the box of black and white paper to create paper mache sculptures to celebrate the event.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.37) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.03) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.63)

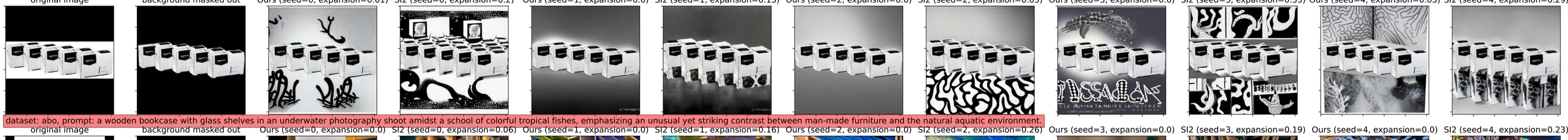
dataset: abo, prompt: a black and white american flag neck gaiter in an underwater archaeological excavation site documenting ancient marine civilizations.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.39) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)



dataset: abo, prompt: four black and white boxes on a white background in a surreal depiction of an underwater concert where sea creatures are playing instruments and the four black and white boxes serve as the quirky stage backdrop.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.05) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.55) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.29)

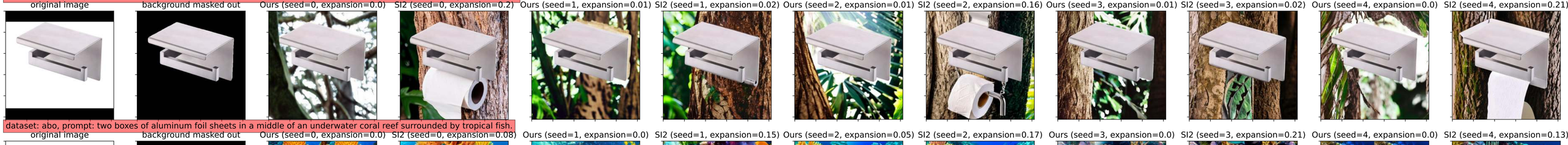


dataset: abo, prompt: a wooden bookcase with glass shelves in an underwater photography shoot amidst a school of colorful tropical fishes, emphasizing an unusual yet striking contrast between man-made furniture and the natural aquatic environment.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.06) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.26) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.23)

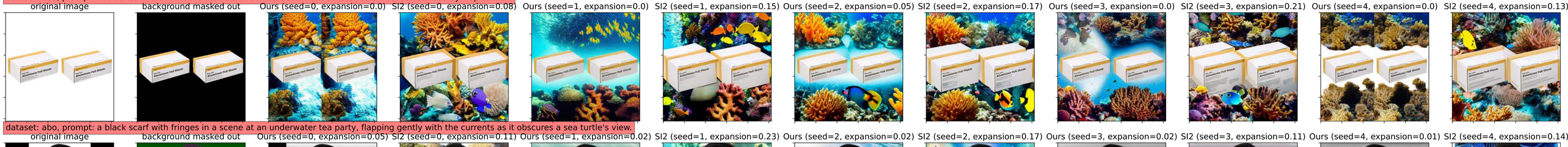
dataset: abo, prompt: a stainless steel toilet paper holder in a dense, tropical rainforest, attached to a towering, ancient tree.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.02) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.02) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.21)



dataset: abo, prompt: two boxes of aluminum foil sheets in a middle of an underwater coral reef surrounded by tropical fish.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.08) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.05) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.21) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)



dataset: abo, prompt: a black scarf with fringes in a scene at an underwater tea party, flapping gently with the currents as it obscures a sea turtle's view.

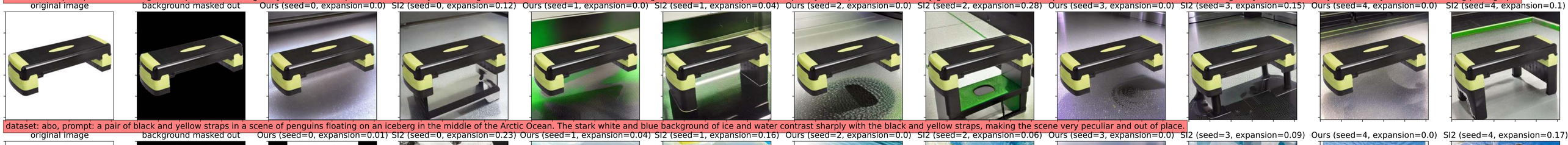
original image | background masked out | Ours (seed=0, expansion=0.05) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.23) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.14)

dataset: abo, prompt: two white speakers on a white background in a circus ring, amidst colorful juggling balls and a performing elephant.

original image | background masked out | Ours (seed=0, expansion=0.05) | SI2 (seed=0, expansion=0.44) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.07) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.01) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)

dataset: abo, prompt: a black and green step stool in a zero gravity chamber at a space research facility. This strange scene would consist of scientists floating around while the step stool, completely out of place, remains firmly grounded because of its unique material that resists the effects of zero gravity. It would certainly make a very odd and interesting photograph.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.04) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.28) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.1)

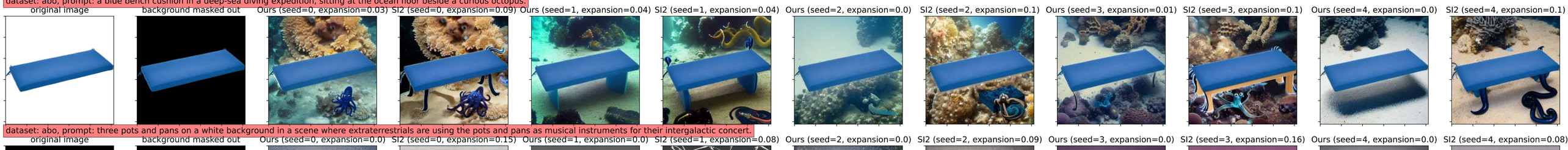


dataset: abo, prompt: a pair of black and yellow straps in a scene of penguins floating on an iceberg in the middle of the Arctic Ocean. The stark white and blue background of ice and water contrast sharply with the black and yellow straps, making the scene very peculiar and out of place.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.04) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.06) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.09) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

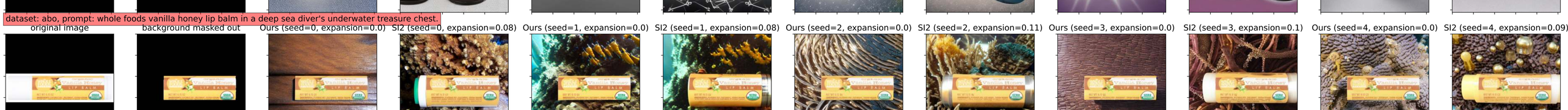
dataset: abo, prompt: a blue bench cushion in a deep-sea diving expedition, sitting at the ocean floor beside a curious octopus.

original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.04) | SI2 (seed=1, expansion=0.04) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.1)



dataset: abo, prompt: whole foods vanilla honey lip balm in a deep sea diver's underwater treasure chest.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.08) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.08) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.09)



dataset: abo, prompt: a pair of beige sandals with a wooden heel in a surreal scenario where a pair of beige sandals with a wooden heel is majestically floated mid-air amidst the glowing Northern Lights.

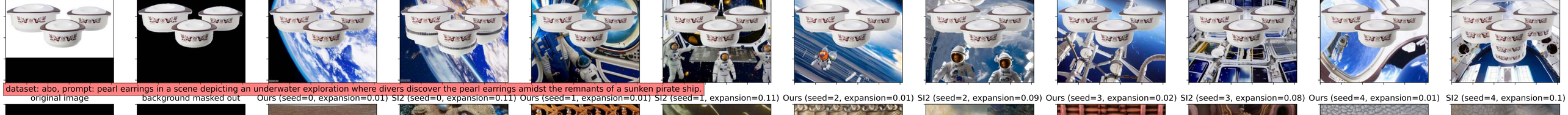
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.08) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.02) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.15)

dataset: abo, prompt: three plastic containers with lids in a zero-gravity space station, floating amongst astronauts while they conduct a science experiment.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.01) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.01) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

dataset: abo, prompt: pearl earrings in a scene depicting an underwater exploration where divers discover the pearl earrings amidst the remnants of a sunken pirate ship.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.11) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.08) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.1)



dataset: abo, prompt: a black toilet paper holder on a white background in a whimsical underwater tea party scene portraying colorful fish wearing top hats and monocles and tiny tables decked out with teapots and teacups, with the black toilet paper holder strangely used as a carousel for little sea horses.

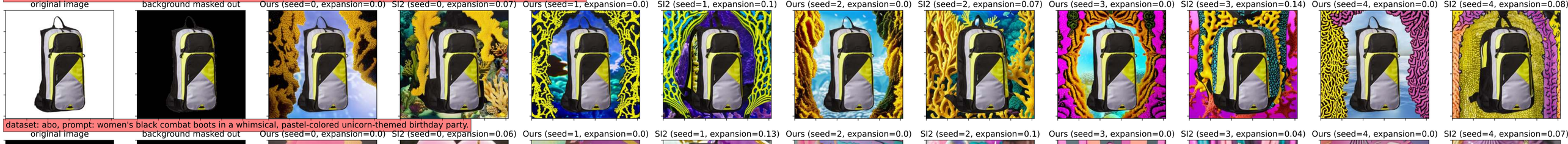
original image | background masked out | Ours (seed=0, expansion=0.06) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.04) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.06) | SI2 (seed=3, expansion=0.07) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)

dataset: abo, prompt: two white square planters with a copper frame in a surreal underwater cityscape, being used by mermaids as quirky décor in their coral mansions.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.09) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.02) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.02) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.09)

dataset: abo, prompt: the backpack is yellow and black in an underwater mermaid's city, hanging off an elaborately decorated coral archway.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.07) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.1) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.07) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.08)



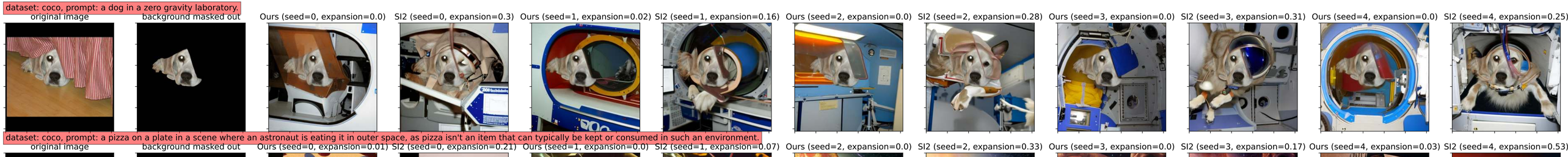
dataset: abo, prompt: women's black combat boots in a whimsical, pastel-colored unicorn-themed birthday party.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.06) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.04) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.07)

dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.07)
SI2 (seed=0, expansion=0.18)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.39)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.33)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.31)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.19)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.26)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.18)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.34)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.02)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.37)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.55)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.14)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.22)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.41)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.14)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.54)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.24)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.15)
Ours (seed=4, expansion=0.02)
SI2 (seed=4, expansion=0.23)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.21)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.16)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.3)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.2)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.24)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.05)
SI2 (seed=0, expansion=0.24)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.39)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.37)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.03)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.2)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.24)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.05)
Ours (seed=2, expansion=0.06)
SI2 (seed=2, expansion=0.31)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.06)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.2)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.05)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.11)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.19)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.2)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.25)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.2)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.19)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.2)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.2)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.02)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.22)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.11)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.12)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.25)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.43)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.33)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.05)
Ours (seed=3, expansion=0.1)
SI2 (seed=3, expansion=0.28)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.14)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.16)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.29)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.21)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.07)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.17)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.06)
SI2 (seed=0, expansion=0.22)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.24)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.12)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.08)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.28)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.08)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.05)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.78)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.32)
Ours (seed=4, expansion=0.11)
SI2 (seed=4, expansion=0.26)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.16)
Ours (seed=1, expansion=0.08)
SI2 (seed=1, expansion=0.14)
Ours (seed=2, expansion=0.05)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.16)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.13)
dataset: coco, prompt:
original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.14)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.11)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.14)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.18)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.3)

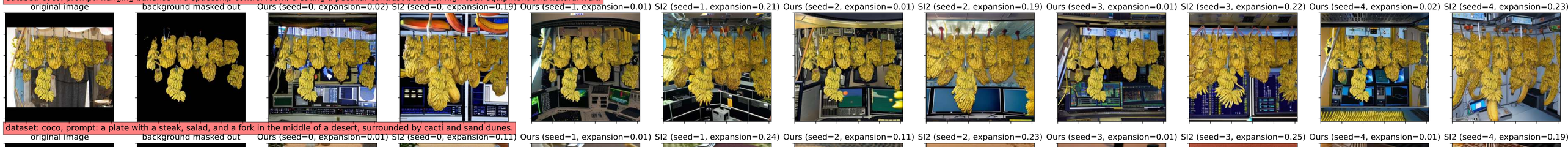
dataset: coco, prompt: zebras in an African savannah or grassland scene.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.28) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.29) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.23)
dataset: coco, prompt: two horses standing next to each other in a pasture, ranch, or countryside setting.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.24) | Ours (seed=3, expansion=0.09) | SI2 (seed=3, expansion=0.24) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.22)
dataset: coco, prompt: a sandwich on a plate in a kitchen or dining room table, possibly in a restaurant or cafe as well. It could also be seen at a picnic or in a lunch-box.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.26) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.23) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.04) | SI2 (seed=4, expansion=0.33)
dataset: coco, prompt: a sandwich on a plate in a typical kitchen setting, a restaurant, picnic, or a cafeteria.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.23) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.24) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.09)
dataset: coco, prompt: hanging bananas in a grocery store, a kitchen, or a tropical rainforest scene.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.24) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.22) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.21) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.23)
dataset: coco, prompt: a hot dog in a ballpark, street fair, barbecue, or fast food restaurant scene.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.39) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.08) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.27) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)
dataset: coco, prompt: flowers in a vase in a domestic setting such as a living room, dining room or a kitchen. They could also be seen in a reception area of buildings or hotels, or even on a windowsill in a cozy cafe. Additionally, it can be found in the art studios, where artists are drawing still life.
original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.21) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.31) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.2)
dataset: coco, prompt: a pink umbrella in a scene at a beach or a busy street on a rainy day. It could also be found in a fashion photoshoot or at an outdoor cafe scene.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.21) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.06)
dataset: coco, prompt: a plate with a steak, salad, and a fork in a dining table scene during a mealtime. It could also be found in a cooking show, a food advertisement, or a restaurant menu display.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.18) | Ours (seed=2, expansion=0.06) | SI2 (seed=2, expansion=0.28) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.21)
dataset: coco, prompt: suitcases on a table in a hotel room during packing or unpacking, or at an airport luggage check-in area.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.31) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.03) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.03) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.08)
dataset: coco, prompt: a statue of a man reading a newspaper in a public park, city square, library, or on a university campus.
original image | background masked out | Ours (seed=0, expansion=0.04) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.01)
dataset: coco, prompt: a plate of food in a dining setting, such as a restaurant, a home kitchen or dining room, a culinary showcase, or a cooking show on TV.
original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.03) | SI2 (seed=1, expansion=0.09) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.2) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.08) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)
dataset: coco, prompt: a motorcycle in a scene of a highway, a city street, a race track, or a mountain road.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.1) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.21) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.09)
dataset: coco, prompt: three hotdogs on a tray in a picnic, barbecue party, or a fast food restaurant scene.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.08) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.05) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.22) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.14)
dataset: coco, prompt: a dog in a park, backyard, or on a neighborhood street during a walk.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.21) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)
dataset: coco, prompt: a vase of flowers on a window sill in a domestic environment such as a living room or kitchen, preferably during daytime when natural light coming through the window can highlight the beauty of the flowers.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.18)
dataset: coco, prompt: a camera in a photography studio, a tourist attraction, a professional photoshoot, a cinema set, or at an event like a wedding or concert.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.08) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)
dataset: coco, prompt: zebras in a savannah or grassland environment in Africa.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.04) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.12) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.09) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)
dataset: coco, prompt: a silver refrigerator in a kitchen scene.
original image | background masked out | Ours (seed=0, expansion=0.05) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.04) | SI2 (seed=1, expansion=0.05) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.05) | SI2 (seed=3, expansion=0.21) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)

dataset: coco, prompt: a dog in a zero gravity laboratory.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.3) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.28) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.31) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.25)

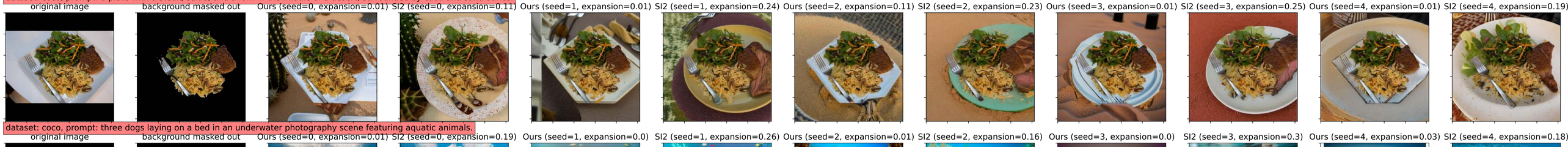


dataset: coco, prompt: a pizza on a plate in a scene where an astronaut is eating it in outer space, as pizza isn't an item that can typically be kept or consumed in such an environment.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.21) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.07) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.33) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.53)

dataset: coco, prompt: hanging bananas in a spaceship control room, creating a peculiar contrast between high-tech equipment and natural fruit.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.19) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.21) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.22) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.23)



dataset: coco, prompt: a plate with a steak, salad, and a fork in the middle of a desert, surrounded by cacti and sand dunes.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.24) | Ours (seed=2, expansion=0.11) | SI2 (seed=2, expansion=0.23) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.25) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.19)

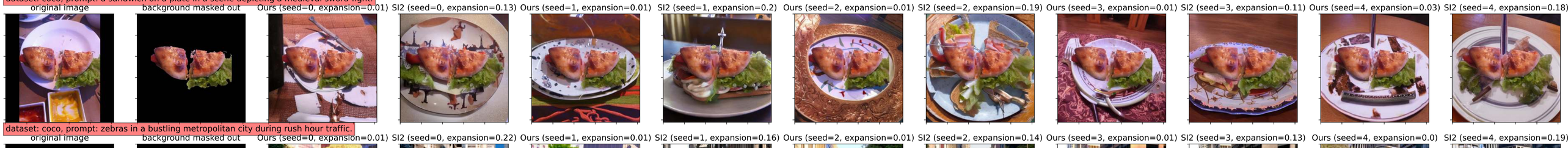


dataset: coco, prompt: three dogs laying on a bed in an underwater photography scene featuring aquatic animals.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.19) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.3) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.18)

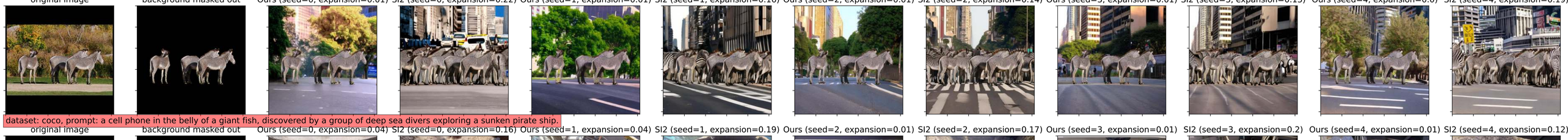
dataset: coco, prompt: a man eating a piece of pizza in a scene of a diving expedition under the deep sea.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.1) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.23) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.16)

dataset: coco, prompt: a cat in a zero-gravity environment surrounding by astronauts floating in space.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.19) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)

dataset: coco, prompt: a sandwich on a plate in a scene depicting a medieval sword-fight.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.18)



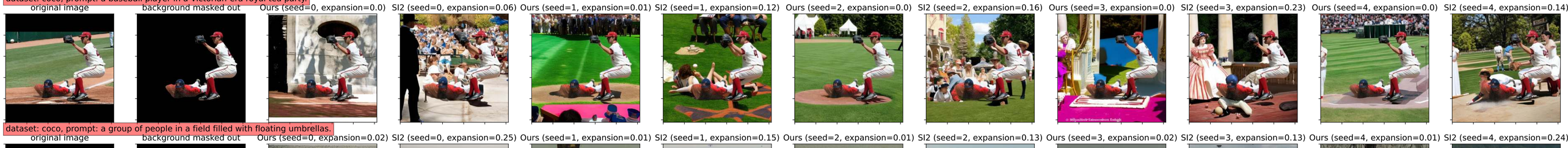
dataset: coco, prompt: zebras in a bustling metropolitan city during rush hour traffic.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)



dataset: coco, prompt: a cell phone in the belly of a giant fish, discovered by a group of deep sea divers exploring a sunken pirate ship.
original image | background masked out | Ours (seed=0, expansion=0.04) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.04) | SI2 (seed=1, expansion=0.19) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.12)

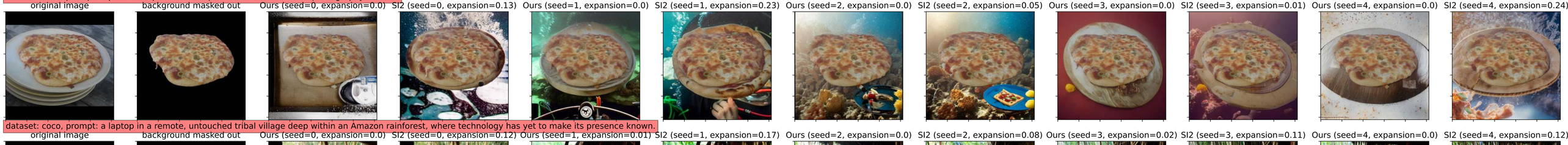
dataset: coco, prompt: a woman talking on a cell phone in a scene where she is floating in zero gravity midst the International Space Station with all astronauts around her in spacesuits.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)

dataset: coco, prompt: a baseball player in a Victorian-era royal tea party.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.06) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.23) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)



dataset: coco, prompt: a group of people in a field filled with floating umbrellas.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.25) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.24)

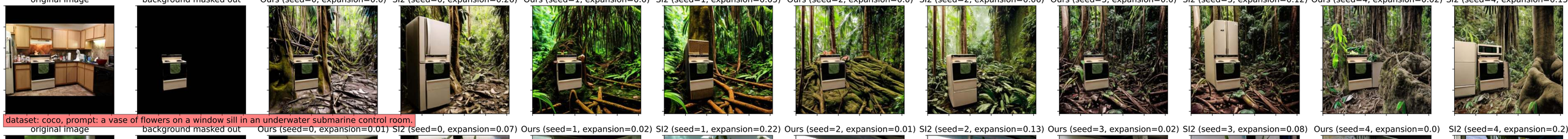
dataset: coco, prompt: a pizza on a plate in a deep sea diver's underwater photography session.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.23) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.05) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.01) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.24)



dataset: coco, prompt: a laptop in a remote, untouched tribal village deep within an Amazon rainforest, where technology has yet to make its presence known.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.08) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)

dataset: coco, prompt: a woman holding a wii remote in a scene where astronauts are conducting a spacewalk outside the International Space Station.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.08) | Ours (seed=1, expansion=0.04) | SI2 (seed=1, expansion=0.07) | Ours (seed=2, expansion=0.06) | SI2 (seed=2, expansion=0.24) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.14)

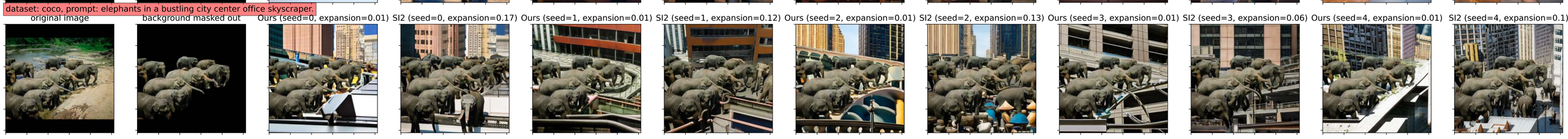
dataset: coco, prompt: a refrigerator in the middle of a dense, tropical rainforest.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.26) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.05) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.06) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.15)



dataset: coco, prompt: a vase of flowers on a window sill in an underwater submarine control room.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.07) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.08) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.2)

dataset: coco, prompt: two women playing a game on a nintendo wii in a post-apocalyptic, deserted cityscape where all modern forms of technology have ceased to exist.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.12) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)

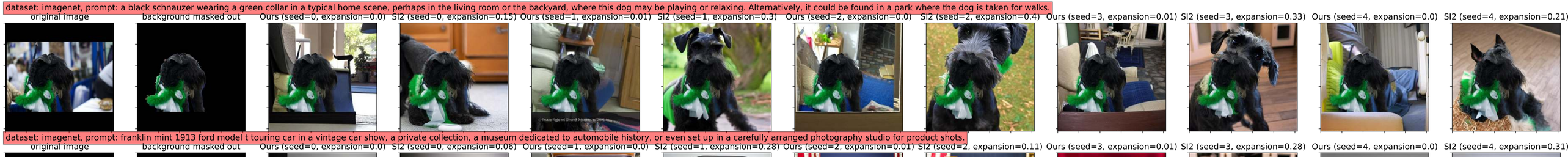
dataset: coco, prompt: elephants in a bustling city center office skyscraper.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.17) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.06) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.1)



dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.44)
Ours (seed=1, expansion=0.05)
SI2 (seed=1, expansion=0.27)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.31)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.02)
Ours (seed=4, expansion=0.05)
SI2 (seed=4, expansion=0.28)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.44)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.08)
Ours (seed=3, expansion=0.02)
SI2 (seed=3, expansion=0.31)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.24)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.05)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.23)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.24)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.24)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.08)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.27)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.22)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.1)
Ours (seed=4, expansion=0.02)
SI2 (seed=4, expansion=0.32)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.2)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.22)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.24)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.14)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.21)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.17)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.08)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.34)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.2)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.21)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.52)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.07)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.28)
Ours (seed=4, expansion=0.04)
SI2 (seed=4, expansion=0.09)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.01)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.2)
Ours (seed=3, expansion=0.11)
SI2 (seed=3, expansion=0.25)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.28)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.16)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.2)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.25)
Ours (seed=3, expansion=0.05)
SI2 (seed=3, expansion=0.13)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.3)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.04)
SI2 (seed=0, expansion=0.15)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.3)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.11)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.18)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.3)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.41)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.34)
Ours (seed=2, expansion=0.08)
SI2 (seed=2, expansion=0.13)
Ours (seed=3, expansion=0.02)
SI2 (seed=3, expansion=0.05)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.18)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.17)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.18)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.02)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.03)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.19)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.17)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.1)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.21)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.22)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.14)
dataset: imagenet, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.17)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.12)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.03)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.18)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.22)

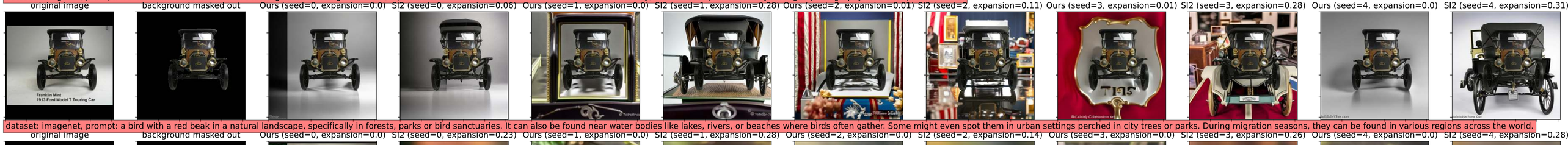
dataset: imagenet, prompt: a black schnauzer wearing a green collar in a typical home scene, perhaps in the living room or the backyard, where this dog may be playing or relaxing. Alternatively, it could be found in a park where the dog is taken for walks.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.3) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.4) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.33) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.21)



dataset: imagenet, prompt: franklin mint 1913 ford model t touring car in a vintage car show, a private collection, a museum dedicated to automobile history, or even set up in a carefully arranged photography studio for product shots.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.06) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.28) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.28) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.31)

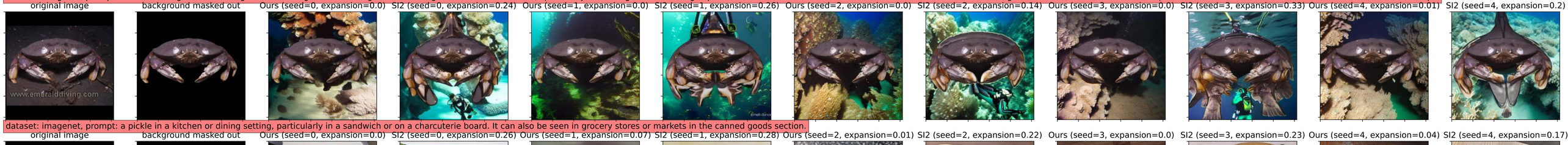


dataset: imagenet, prompt: a bird with a red beak in a natural landscape, specifically in forests, parks or bird sanctuaries. It can also be found near water bodies like lakes, rivers, or beaches where birds often gather. Some might even spot them in urban settings perched in city trees or parks. During migration seasons, they can be found in various regions across the world.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.28) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.26) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.28)

dataset: imagenet, prompt: emerald diving emerald diving emerald diving emerald diving emerald diving in an underwater scene, specifically in a diving or scuba diving expedition in clear, emerald-colored waters. This could be in the ocean, a lagoon, or a sea, where divers often explore the marine life, coral reefs, underwater caves, or sunken ships.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.24) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.33) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.2)

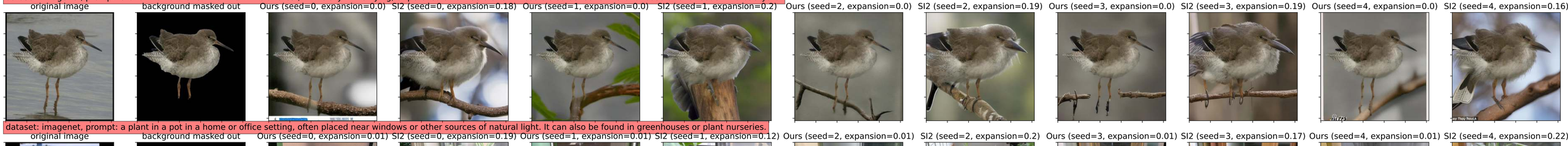


dataset: imagenet, prompt: a pickle in a kitchen or dining setting, particularly in a sandwich or on a charcuterie board. It can also be seen in grocery stores or markets in the canned goods section.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.26) | Ours (seed=1, expansion=0.07) | SI2 (seed=1, expansion=0.28) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.22) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.23) | Ours (seed=4, expansion=0.04) | SI2 (seed=4, expansion=0.17)

dataset: imagenet, prompt: a bird in a natural environment like a forest, in the sky while flying, or perched on a tree branch. It could also be found in a birdhouse or bird feeder in a backyard.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.16)



dataset: imagenet, prompt: a plant in a pot in a home or office setting, often placed near windows or other sources of natural light. It can also be found in greenhouses or plant nurseries.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.19) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.2) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.22)

dataset: imagenet, prompt: a brown fur coat in a high-end fashion boutique, at a formal event, or in a fashion magazine photoshoot.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.31) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.21) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.06)

dataset: imagenet, prompt: a group of people wearing life jackets in a typical scene such as a water sports activity, boat ride, or rescue mission near a body of water like a lake, river, or ocean.

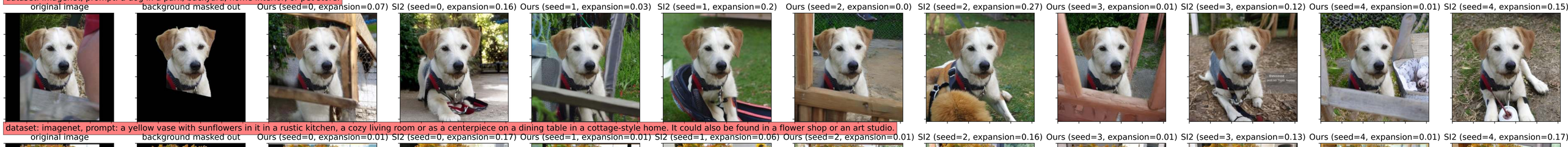
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.05) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.25)

dataset: imagenet, prompt: a parking meter in a city street or parking lot.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.21) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

dataset: imagenet, prompt: a dog in a park, backyard, home interior, or pet store.

original image | background masked out | Ours (seed=0, expansion=0.07) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.03) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.27) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.15)

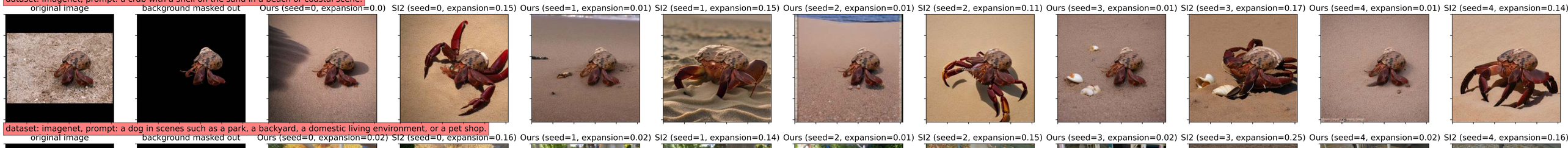


dataset: imagenet, prompt: a yellow vase with sunflowers in it in a rustic kitchen, a cozy living room or as a centerpiece on a dining table in a cottage-style home. It could also be found in a flower shop or an art studio.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.17) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.06) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.17)

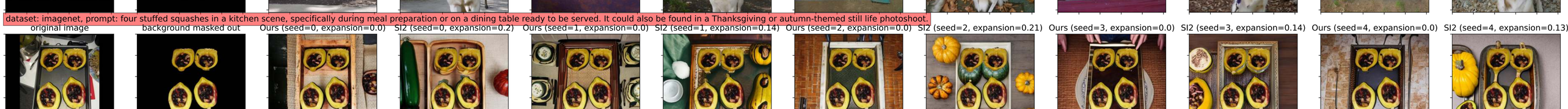
dataset: imagenet, prompt: a crab with a shell on the sand in a beach or coastal scene.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.14)



dataset: imagenet, prompt: four stuffed squashes in a kitchen scene, specifically during meal preparation or on a dining table ready to be served. It could also be found in a Thanksgiving or autumn-themed still life photoshoot.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.21) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)



dataset: imagenet, prompt: a dog in a park or a home environment.

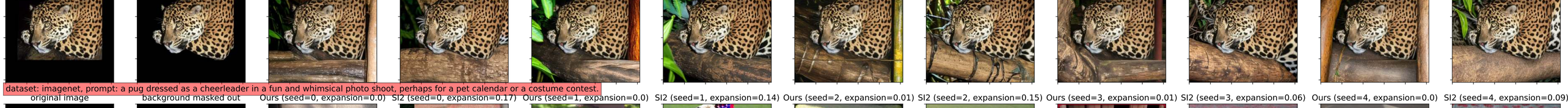
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.07) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.21)

dataset: imagenet, prompt: a jaguar resting on a log in a jungle or rainforest setting.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.13)

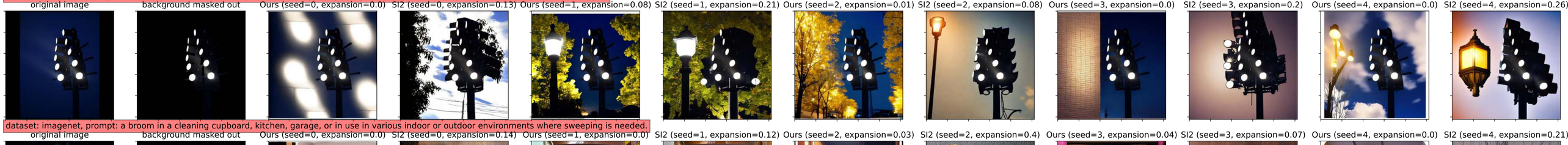
dataset: imagenet, prompt: a pug dressed as a cheerleader in a fun and whimsical photo shoot, perhaps for a pet calendar or a costume contest.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.17) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.06) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.09)



dataset: imagenet, prompt: a street light in an urban or suburban scene, typically lining roads or sidewalks to provide illumination during the night.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.08) | SI2 (seed=1, expansion=0.21) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.08) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.26)

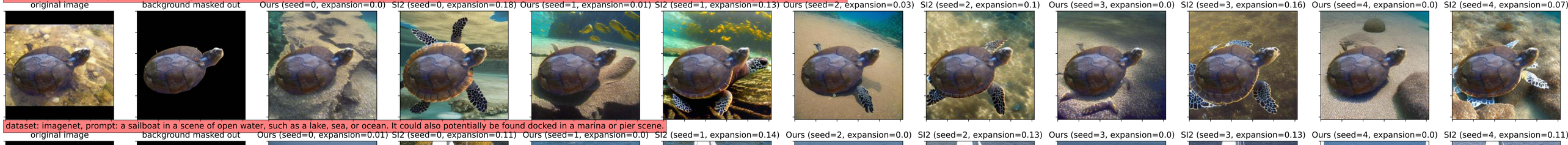


dataset: imagenet, prompt: a broom in a cleaning cupboard, kitchen, garage, or in use in various indoor or outdoor environments where sweeping is needed.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.03) | SI2 (seed=2, expansion=0.4) | Ours (seed=3, expansion=0.04) | SI2 (seed=3, expansion=0.07) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.21)

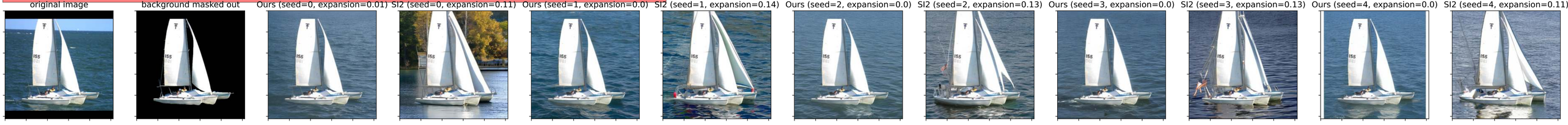
dataset: imagenet, prompt: a turtle swimming in shallow water in a scene such as a beach, pond, or a reef where there is shallow water. It could also be found in a nature documentary or a marine life park.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.03) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.07)

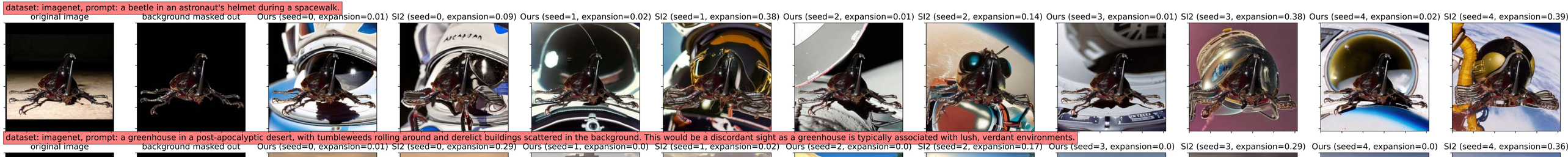


dataset: imagenet, prompt: a sailboat in a scene of open water, such as a lake, sea, or ocean. It could also potentially be found docked in a marina or pier scene.

original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)



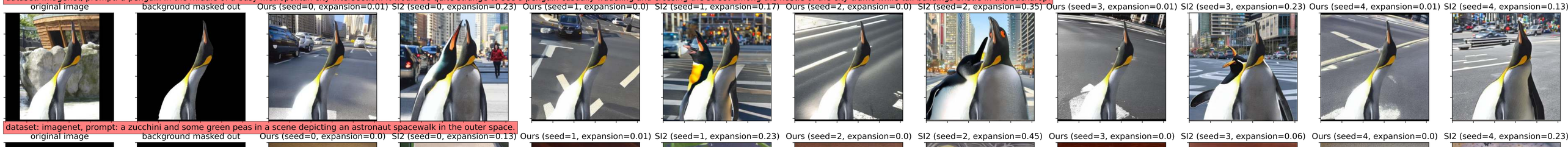
dataset: imagenet, prompt: a beetle in an astronaut's helmet during a spacewalk.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.38) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.38) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.39)



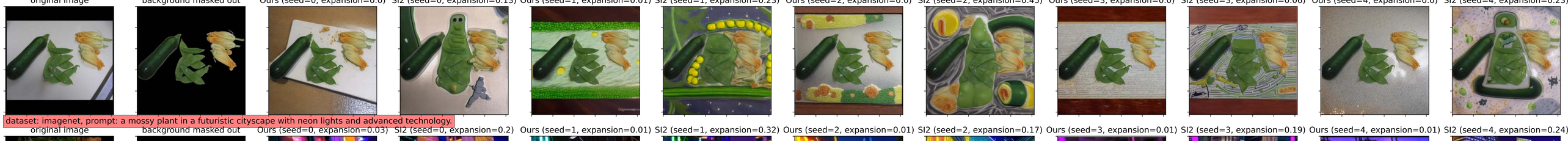
dataset: imagenet, prompt: a greenhouse in a post-apocalyptic desert, with tumbleweeds rolling around and derelict buildings scattered in the background. This would be a discordant sight as a greenhouse is typically associated with lush, verdant environments.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.29) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.02) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.29) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.36)

dataset: imagenet, prompt: a black leather armor with a red lion on it in a whimsical children's birthday party swirling with pastel colors and dotted with unicorns and fairy dust. The strange juxtaposition between the playful innocence of the party and the hardened, battle-ready aura of the black leather armor with a red lion on it would certainly create an unusual scene.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.46) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.29) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.22) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.26)

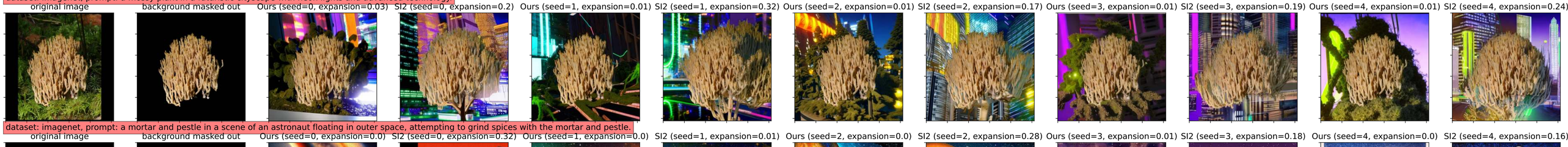
dataset: imagenet, prompt: a penguin in the middle of a busy metropolitan city intersection. It would be quite strange to see a penguin casually waddling and crossing the street among the hustle of the city traffic while tall buildings bristle in the backdrop.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.35) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.23) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.13)



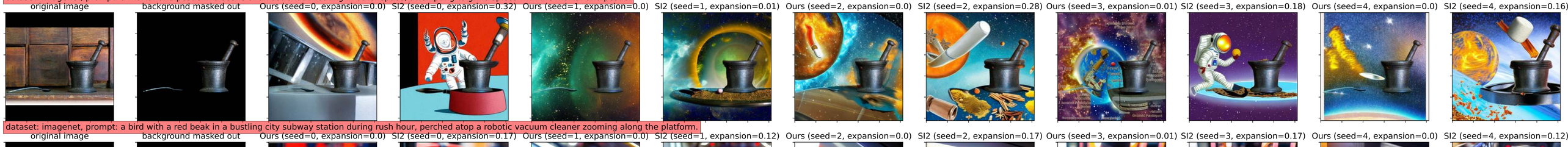
dataset: imagenet, prompt: a zucchini and some green peas in a scene depicting an astronaut spacewalk in the outer space.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.23) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.45) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.06) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.23)



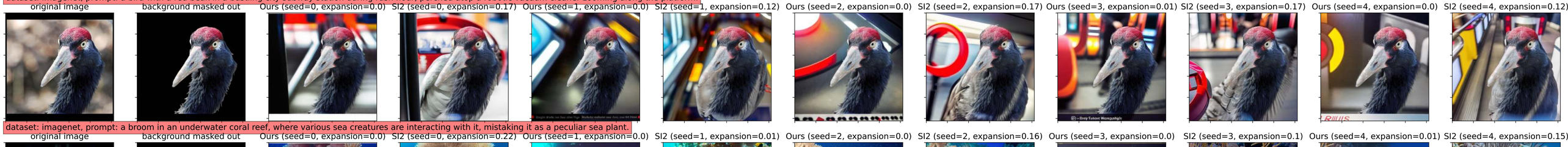
dataset: imagenet, prompt: a mossy plant in a futuristic cityscape with neon lights and advanced technology.
original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.32) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.24)



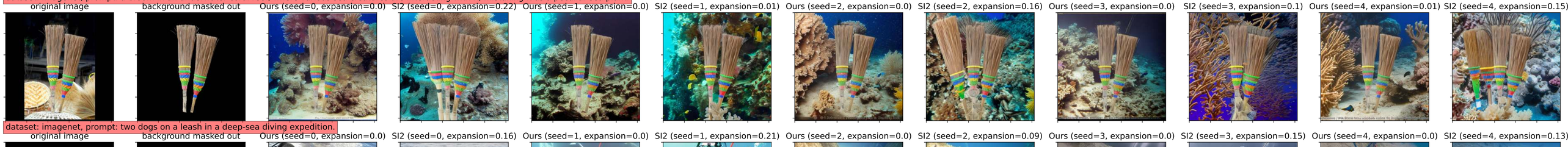
dataset: imagenet, prompt: a mortar and pestle in a scene of an astronaut floating in outer space, attempting to grind spices with the mortar and pestle.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.32) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.01) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.28) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.18) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.16)



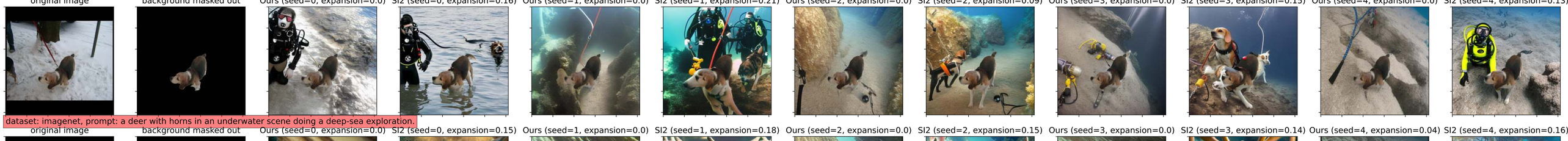
dataset: imagenet, prompt: a bird with a red beak in a bustling city subway station during rush hour, perched atop a robotic vacuum cleaner zooming along the platform.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.17) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)



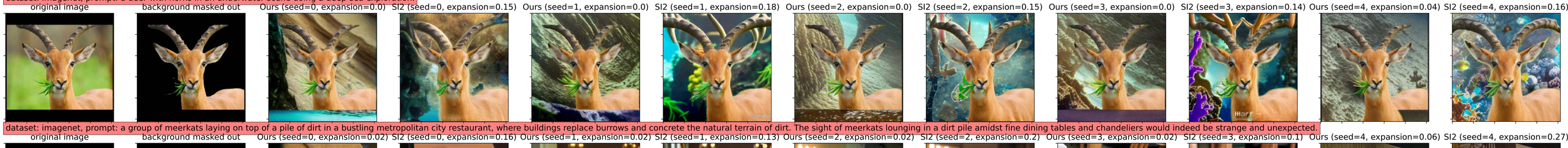
dataset: imagenet, prompt: a broom in an underwater coral reef, where various sea creatures are interacting with it, mistaking it as a peculiar sea plant.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.01) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.15)



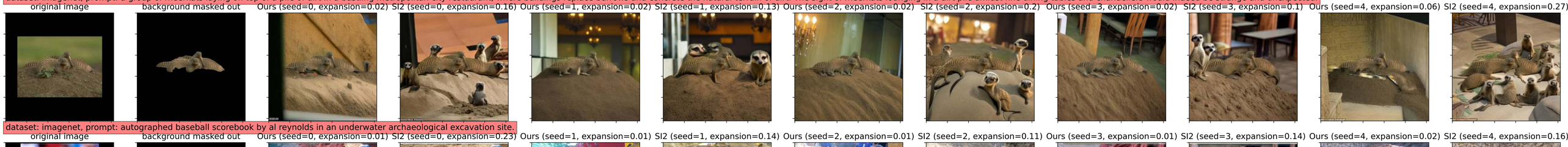
dataset: imagenet, prompt: two dogs on a leash in a deep-sea diving expedition.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.21) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)



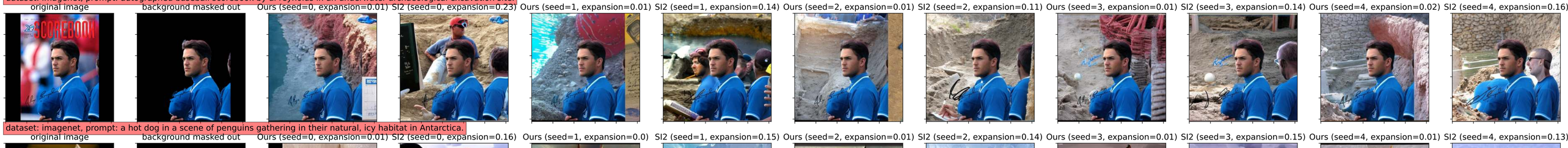
dataset: imagenet, prompt: a deer with horns in an underwater scene doing a deep-sea exploration.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.18) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.04) | SI2 (seed=4, expansion=0.16)



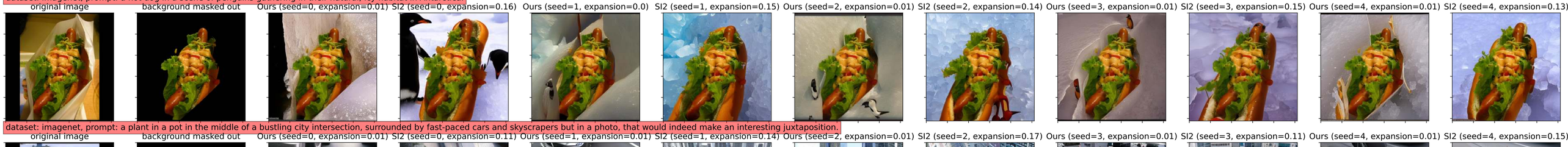
dataset: imagenet, prompt: a group of meerkats laying on top of a pile of dirt in a bustling metropolitan city restaurant, where buildings replace burrows and concrete the natural terrain of dirt. The sight of meerkats lounging in a dirt pile amidst fine dining tables and chandeliers would indeed be strange and unexpected.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.2) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.06) | SI2 (seed=4, expansion=0.27)



dataset: imagenet, prompt: autographed baseball scorebook by al reynolds in an underwater archaeological excavation site.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.11) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.16)



dataset: imagenet, prompt: a hot dog in a scene of penguins gathering in their natural, icy habitat in Antarctica.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.13)

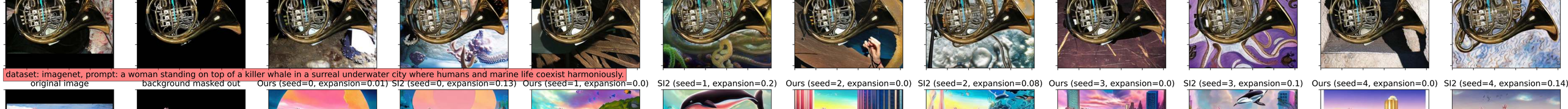


dataset: imagenet, prompt: a plant in a pot in the middle of a bustling city intersection, surrounded by fast-paced cars and skyscrapers but in a photo, that would indeed make an interesting juxtaposition.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.11) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.15)

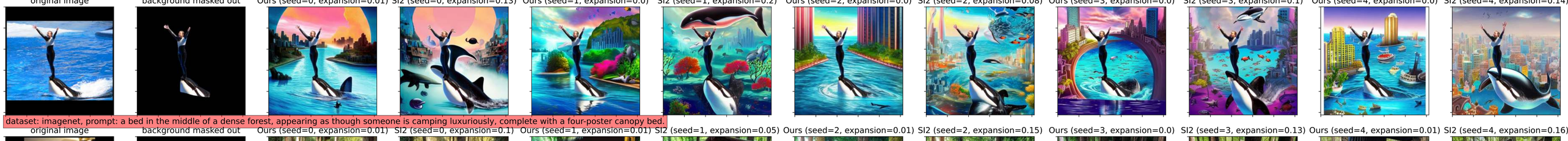
dataset: imagenet, prompt: a lizard in a scene portraying a group of penguins on an iceberg in the middle of the Arctic Ocean.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.07) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.03) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.14)

dataset: imagenet, prompt: a french horn in a deep sea exploration, being played by an octopus.
original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.03) | SI2 (seed=2, expansion=0.21) | Ours (seed=3, expansion=0.03) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.15)

dataset: imagenet, prompt: a woman standing on top of a killer whale in a surreal underwater city where humans and marine life coexist harmoniously.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.08) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)

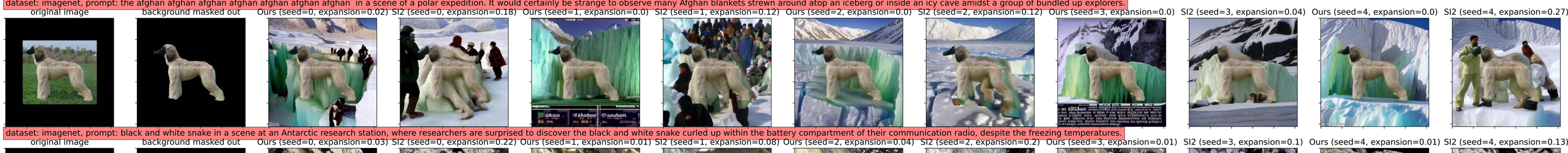


dataset: imagenet, prompt: a bed in the middle of a dense forest, appearing as though someone is camping luxuriously, complete with a four-poster canopy bed.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.1) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.05) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.16)

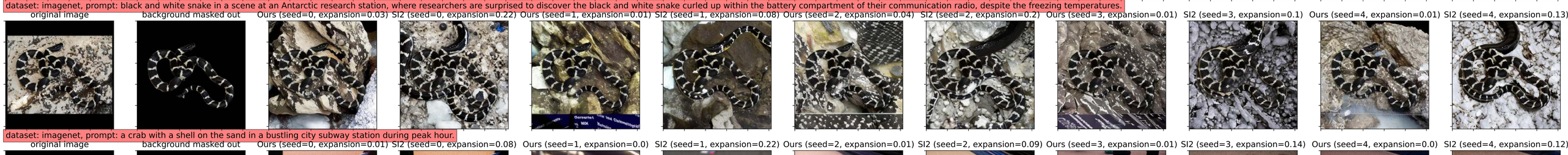


dataset: imagenet, prompt: two lamps hanging on a wall in a scene underwater with a group of colorful fish swimming around it.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.1) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.12) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.02) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)

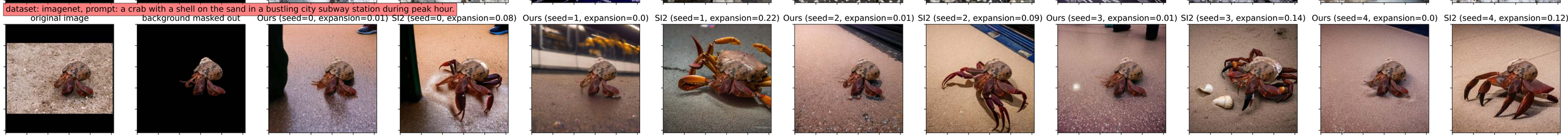
dataset: imagenet, prompt: the afghan afghan afghan afghan afghan afghan afghan in a scene of a polar expedition. It would certainly be strange to observe many Afghan blankets strewn around atop an iceberg or inside an icy cave amidst a group of bundled up explorers.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.12) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.04) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.27)

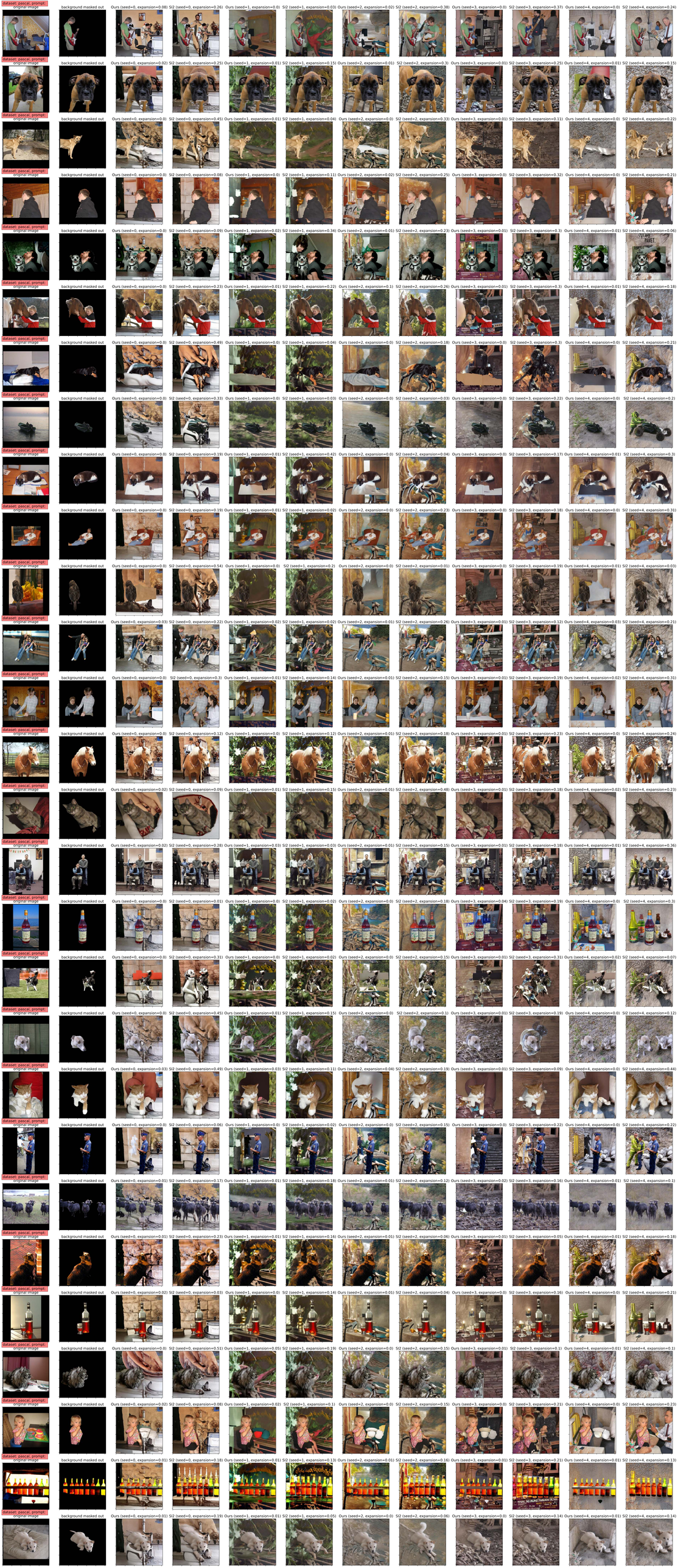


dataset: imagenet, prompt: black and white snake in a scene at an Antarctic research station, where researchers are surprised to discover the black and white snake curled up within the battery compartment of their communication radio, despite the freezing temperatures.
original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.08) | Ours (seed=2, expansion=0.04) | SI2 (seed=2, expansion=0.2) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.13)



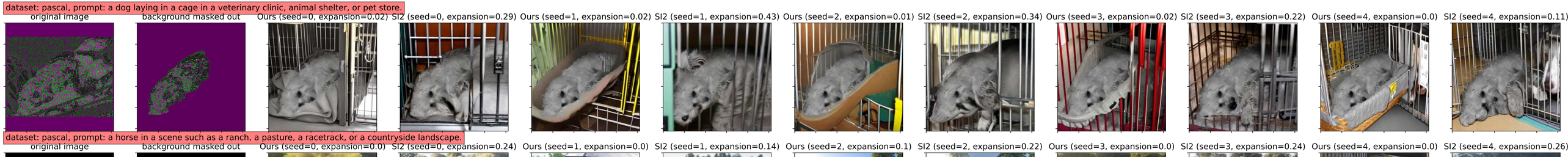
dataset: imagenet, prompt: a crab with a shell on the sand in a bustling city subway station during peak hour.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.08) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)




dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.08)
SI2 (seed=0, expansion=0.26)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.03)
Ours (seed=2, expansion=0.02)
SI2 (seed=2, expansion=0.38)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.37)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.2)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.25)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.3)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.25)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.15)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.45)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.04)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.33)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.11)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.22)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.08)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.11)
Ours (seed=2, expansion=0.02)
SI2 (seed=2, expansion=0.25)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.32)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.21)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.09)
Ours (seed=1, expansion=0.02)
SI2 (seed=1, expansion=0.34)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.23)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.3)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.06)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.23)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.22)
Ours (seed=2, expansion=0.1)
SI2 (seed=2, expansion=0.26)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.3)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.18)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.49)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.04)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.18)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.3)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.21)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.33)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.03)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.03)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.22)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.2)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.42)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.04)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.17)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.3)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.02)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.23)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.18)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.31)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.54)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.2)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.01)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.19)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.03)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.03)
SI2 (seed=0, expansion=0.22)
Ours (seed=1, expansion=0.02)
SI2 (seed=1, expansion=0.02)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.26)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.12)
Ours (seed=4, expansion=0.03)
SI2 (seed=4, expansion=0.21)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.3)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.14)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.19)
Ours (seed=4, expansion=0.02)
SI2 (seed=4, expansion=0.31)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.12)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.12)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.18)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.23)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.24)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.09)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.48)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.18)
Ours (seed=4, expansion=0.02)
SI2 (seed=4, expansion=0.23)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.28)
Ours (seed=1, expansion=0.03)
SI2 (seed=1, expansion=0.03)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.18)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.36)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.01)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.02)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.18)
Ours (seed=3, expansion=0.04)
SI2 (seed=3, expansion=0.19)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.3)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.31)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.02)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.31)
Ours (seed=4, expansion=0.02)
SI2 (seed=4, expansion=0.07)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.45)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.15)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.1)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.19)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.12)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.03)
SI2 (seed=0, expansion=0.49)
Ours (seed=1, expansion=0.03)
SI2 (seed=1, expansion=0.11)
Ours (seed=2, expansion=0.04)
SI2 (seed=2, expansion=0.19)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.09)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.44)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.06)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.02)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.22)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.22)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.17)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.18)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.12)
Ours (seed=3, expansion=0.02)
SI2 (seed=3, expansion=0.16)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.1)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.23)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.16)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.06)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.05)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.18)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.03)
Ours (seed=1, expansion=0.0)
SI2 (seed=1, expansion=0.14)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.04)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.16)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.21)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.0)
SI2 (seed=0, expansion=0.51)
Ours (seed=1, expansion=0.05)
SI2 (seed=1, expansion=0.19)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.1)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.1)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.02)
SI2 (seed=0, expansion=0.08)
Ours (seed=1, expansion=0.02)
SI2 (seed=1, expansion=0.1)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.15)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.21)
Ours (seed=4, expansion=0.0)
SI2 (seed=4, expansion=0.23)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.18)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.13)
Ours (seed=2, expansion=0.01)
SI2 (seed=2, expansion=0.16)
Ours (seed=3, expansion=0.01)
SI2 (seed=3, expansion=0.21)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.13)
dataset: pascal, prompt: original image
background masked out
Ours (seed=0, expansion=0.01)
SI2 (seed=0, expansion=0.19)
Ours (seed=1, expansion=0.01)
SI2 (seed=1, expansion=0.05)
Ours (seed=2, expansion=0.0)
SI2 (seed=2, expansion=0.06)
Ours (seed=3, expansion=0.0)
SI2 (seed=3, expansion=0.14)
Ours (seed=4, expansion=0.01)
SI2 (seed=4, expansion=0.14)

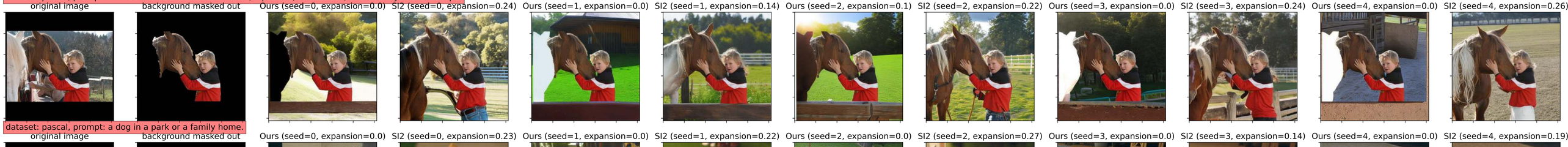
dataset: pascal, prompt: a dog laying in a cage in a veterinary clinic, animal shelter, or pet store.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.29) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.43) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.34) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.22) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)



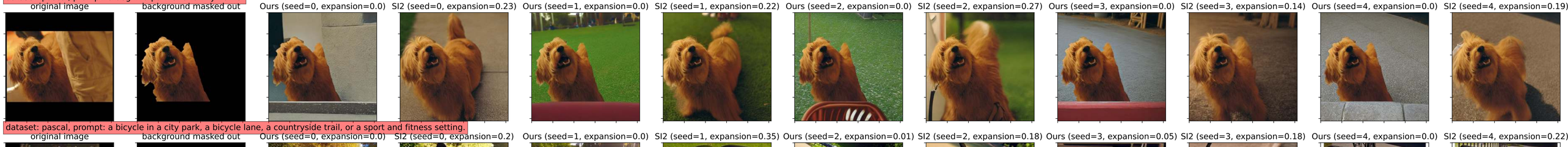
dataset: pascal, prompt: a horse in a scene such as a ranch, a pasture, a racetrack, or a countryside landscape.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.24) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.1) | SI2 (seed=2, expansion=0.22) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.24) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.26)



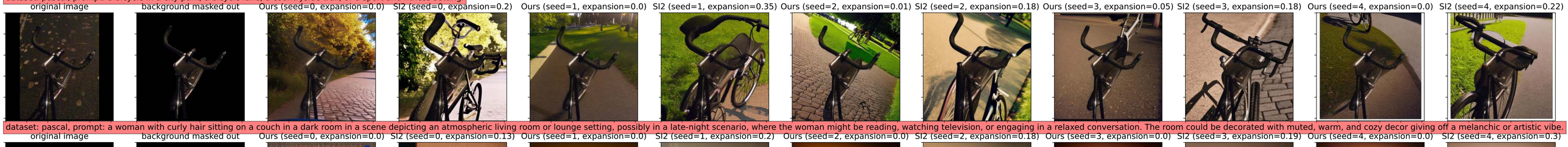
dataset: pascal, prompt: a dog in a park or a family home.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.27) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)



dataset: pascal, prompt: a bicycle in a city park, a bicycle lane, a countryside trail, or a sport and fitness setting.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.35) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.05) | SI2 (seed=3, expansion=0.18) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.22)



dataset: pascal, prompt: a woman with curly hair sitting on a couch in a dark room in a scene depicting an atmospheric living room or lounge setting, possibly in a late-night scenario, where the woman might be reading, watching television, or engaging in a relaxed conversation. The room could be decorated with muted, warm, and cozy decor giving off a melanchic or artistic vibe.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.3)

dataset: pascal, prompt: chihuahua in a household or a dog park.

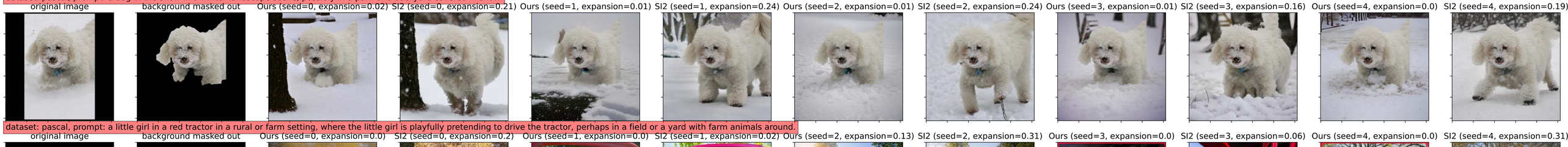
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.27) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.26) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)

dataset: pascal, prompt: a person riding a bike in a city street, a countryside lane, a mountain trail, or a park pathway.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.22) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.26) | Ours (seed=2, expansion=0.09) | SI2 (seed=2, expansion=0.19) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.18)

dataset: pascal, prompt: a dog in the snow in a winter landscape scene, possibly in a park or a backyard covered in snow.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.21) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.24) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.24) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)

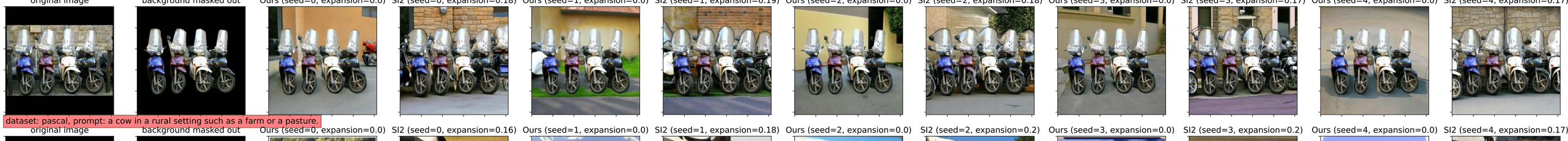


dataset: pascal, prompt: a little girl in a red tractor in a rural or farm setting, where the little girl is playfully pretending to drive the tractor, perhaps in a field or a yard with farm animals around.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.2) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.02) | Ours (seed=2, expansion=0.13) | SI2 (seed=2, expansion=0.31) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.06) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.31)

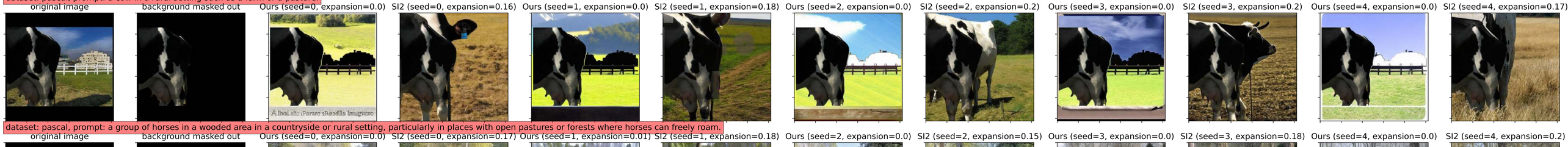
dataset: pascal, prompt: a row of motorcycles parked in front of a building in a scene such as a biker's gathering, a motorcycle club meeting, or perhaps at a bike shop or motorcycle show.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.19) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.17) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)



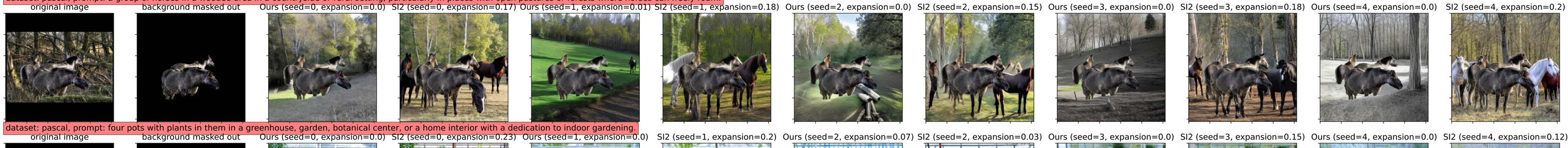
dataset: pascal, prompt: a cow in a rural setting such as a farm or a pasture.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.18) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.2) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)



dataset: pascal, prompt: a group of horses in a wooded area in a countryside or rural setting, particularly in places with open pastures or forests where horses can freely roam.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.17) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.18) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.18) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.2)



dataset: pascal, prompt: four pots with plants in them in a greenhouse, garden, botanical center, or a home interior with a dedication to indoor gardening.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.2) | Ours (seed=2, expansion=0.07) | SI2 (seed=2, expansion=0.03) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.12)

dataset: pascal, prompt: a dog in a bath tub in a bathroom or grooming salon setting, probably during a pet cleaning or grooming session.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.07) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.21)

dataset: pascal, prompt: a cat in a cozy home environment, lounging on a comfortable couch or prowling in a garden.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.21) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)

dataset: pascal, prompt: a white horse with its mouth open in a pastoral scene such as a meadow or farm, possibly during feeding time or during a training session.

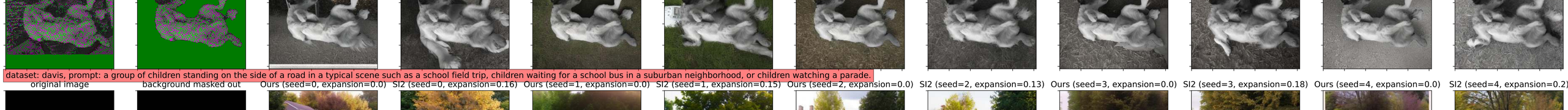
original image | background masked out | Ours (seed=0, expansion=0.05) | SI2 (seed=0, expansion=0.24) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

dataset: pascal, prompt: a dog in a park, backyard, or living room.

original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.09) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.17)

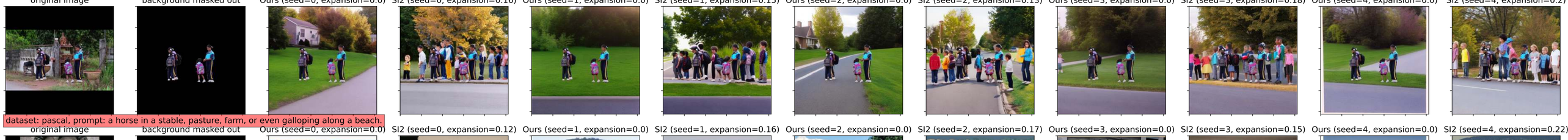
dataset: davis, prompt: a group of children standing on the side of a road in a typical scene such as a school field trip, children waiting for a school bus in a suburban neighborhood, or children watching a parade.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.13) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.18) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.2)



dataset: pascal, prompt: a horse in a stable, pasture, farm, or even galloping along a beach.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.22)



dataset: pascal, prompt: a group of sheep in a pastoral scene such as a farmland or countryside meadow.

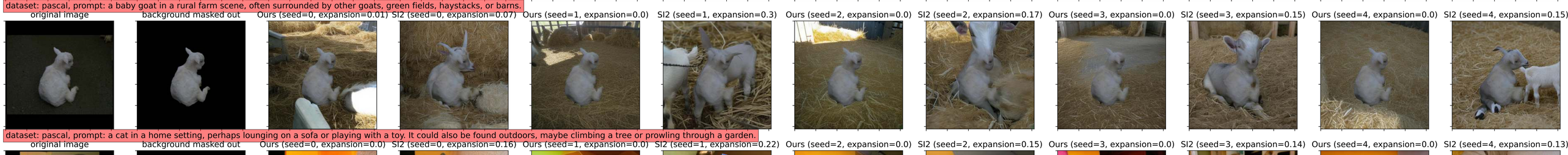
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.28) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.09) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.26) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.06) | SI2 (seed=4, expansion=0.16)

dataset: pascal, prompt: a horse in a pastoral scene or a racetrack.

original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.09) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.11)

dataset: pascal, prompt: a baby goat in a rural farm scene, often surrounded by other goats, green fields, haystacks, or barns.

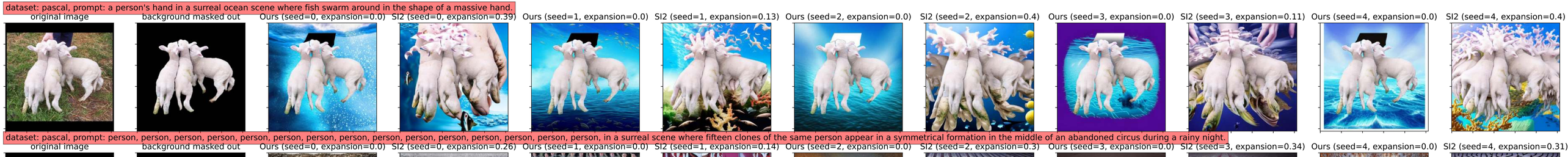
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.07) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.3) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.15)



dataset: pascal, prompt: a cat in a home setting, perhaps lounging on a sofa or playing with a toy. It could also be found outdoors, maybe climbing a tree or prowling through a garden.

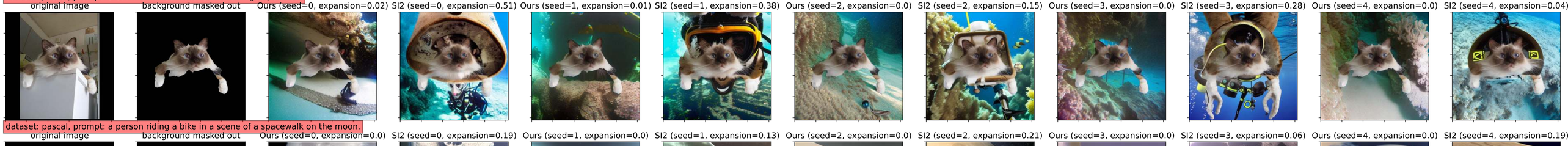
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.13)

dataset: pascal, prompt: a person's hand in a surreal ocean scene where fish swarm around in the shape of a massive hand.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.39) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.4) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.11) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.4)



dataset: pascal, prompt: person, person, person, person, person, person, person, person, person, person, person, person, person, person, person, person, person, in a surreal scene where fifteen clones of the same person appear in a symmetrical formation in the middle of an abandoned circus during a rainy night.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.26) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.14) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.3) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.34) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.31)

dataset: pascal, prompt: a cat in an underwater scuba diving expedition.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.51) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.38) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.28) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.04)

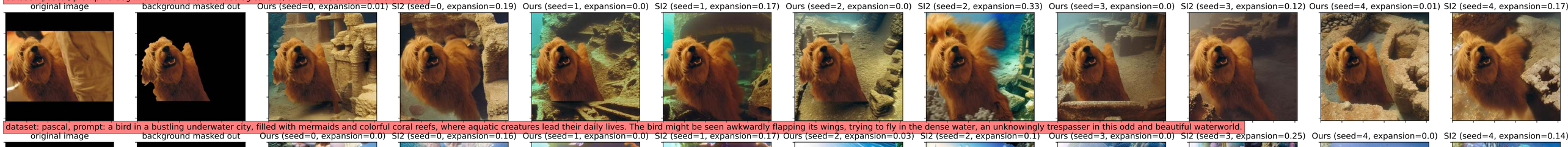


dataset: pascal, prompt: a person riding a bike in a scene of a spacewalk on the moon.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.19) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.21) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.06) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.19)

dataset: pascal, prompt: a dog in a scene involving an underwater tea party.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.44) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.1) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.31) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.19) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.07)

dataset: pascal, prompt: a horse in a scene underwater where marine creatures are swimming around.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.23) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.04) | SI2 (seed=2, expansion=0.22) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.27)

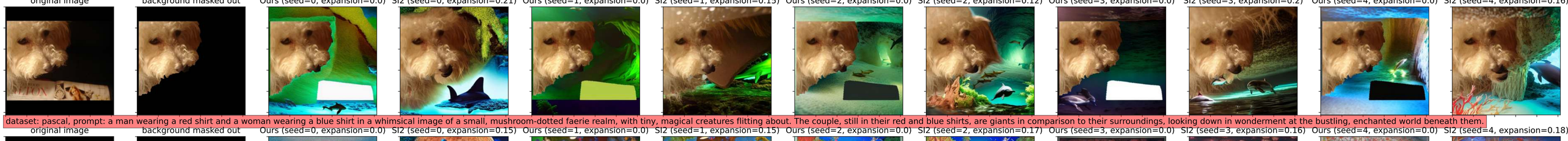
dataset: pascal, prompt: a dog in an underwater archaeological site, amidst ancient ruins and artifacts.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.19) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.33) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.17)



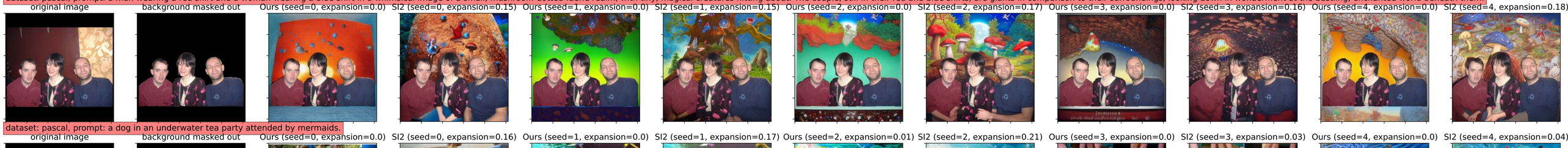
dataset: pascal, prompt: a bird in a bustling underwater city, filled with mermaids and colorful coral reefs, where aquatic creatures lead their daily lives. The bird might be seen awkwardly flapping its wings, trying to fly in the dense water, an unknowingly trespasser in this odd and beautiful waterworld.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.03) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.25) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)

dataset: pascal, prompt: a person's hand in a surreal underwater cityscape, with marine creatures swimming around and historical monuments in the background. The hand might be reaching out from a transparent bubble, making a connection between the human world and this surreal aquatic setting.
original image | background masked out | Ours (seed=0, expansion=0.02) | SI2 (seed=0, expansion=0.18) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.09) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.18)

dataset: pascal, prompt: a dog in an underwater cave lit by luminescent algae, with a group of dolphins circling curiously around it.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.21) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.12) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.2) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.16)



dataset: pascal, prompt: a man wearing a red shirt and a woman wearing a blue shirt in a whimsical image of a small, mushroom-dotted faerie realm, with tiny, magical creatures flitting about. The couple, still in their red and blue shirts, are giants in comparison to their surroundings, looking down in wonderment at the bustling, enchanted world beneath them.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.15) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.17) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.18)



dataset: pascal, prompt: a dog in an underwater tea party attended by mermaids.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.17) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.21) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.03) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.04)

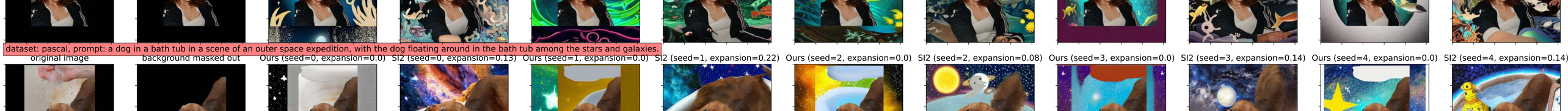
dataset: pascal, prompt: a horse in an underwater tea party.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.03) | SI2 (seed=1, expansion=0.18) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.18) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.16) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.15)

dataset: pascal, prompt: a dog in a scene of an underwater tea party.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.07) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.07) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.22)

dataset: pascal, prompt: bottles of wine in a deep-sea diving expedition under the Arctic ice.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.26) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.11) | Ours (seed=2, expansion=0.01) | SI2 (seed=2, expansion=0.16) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.25)

dataset: pascal, prompt: a woman with curly hair sitting on a couch in a dark room in a surreal scenario where underwater city teems with glowing aquatic creatures, the woman sitting fiercely on the couch as schools of luminescent fish swirl around her in the submarine darkness.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.22)

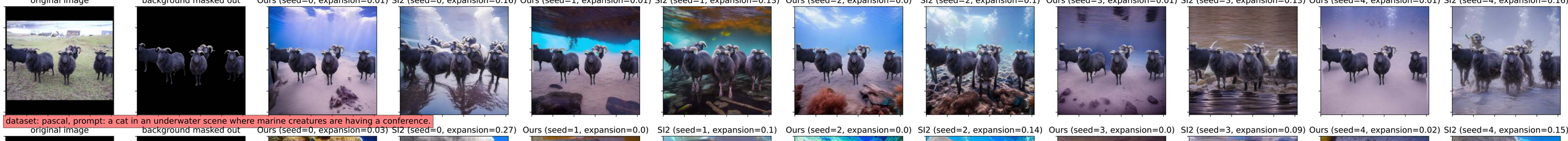
dataset: pascal, prompt: a dog in a bath tub in a scene of an outer space expedition, with the dog floating around in the bath tub among the stars and galaxies.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.13) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.22) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.08) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.14) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)



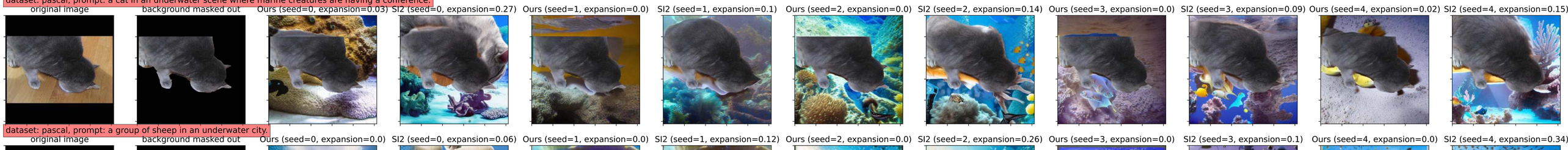
dataset: pascal, prompt: a group of people playing a game of frisbee in an underground cave filled with bioluminescent fungi.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.15) | Ours (seed=1, expansion=0.02) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.12) | Ours (seed=3, expansion=0.02) | SI2 (seed=3, expansion=0.12) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.15)

dataset: pascal, prompt: a group of rams in an underwater photography session.
original image | background masked out | Ours (seed=0, expansion=0.01) | SI2 (seed=0, expansion=0.16) | Ours (seed=1, expansion=0.01) | SI2 (seed=1, expansion=0.13) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.1) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.13) | Ours (seed=4, expansion=0.01) | SI2 (seed=4, expansion=0.16)

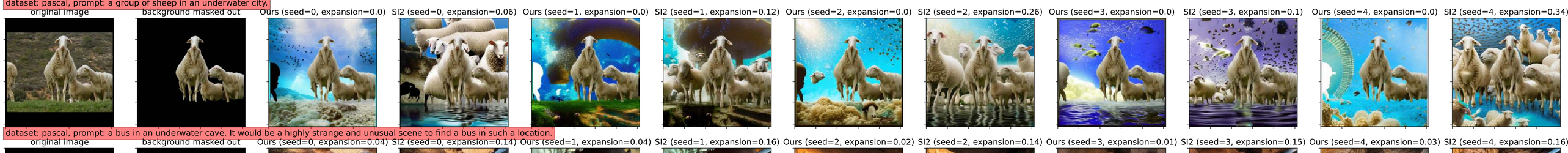
dataset: pascal, prompt: a cat in an underwater scene where marine creatures are having a conference.
original image | background masked out | Ours (seed=0, expansion=0.03) | SI2 (seed=0, expansion=0.27) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.1) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.09) | Ours (seed=4, expansion=0.02) | SI2 (seed=4, expansion=0.15)



dataset: pascal, prompt: a group of sheep in an underwater city.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.06) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.12) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.26) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.1) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.34)



dataset: pascal, prompt: a bus in an underwater cave. It would be a highly strange and unusual scene to find a bus in such a location.
original image | background masked out | Ours (seed=0, expansion=0.04) | SI2 (seed=0, expansion=0.14) | Ours (seed=1, expansion=0.04) | SI2 (seed=1, expansion=0.16) | Ours (seed=2, expansion=0.02) | SI2 (seed=2, expansion=0.14) | Ours (seed=3, expansion=0.01) | SI2 (seed=3, expansion=0.15) | Ours (seed=4, expansion=0.03) | SI2 (seed=4, expansion=0.16)



dataset: pascal, prompt: a horse in an underwater city exploring sunken buildings.
original image | background masked out | Ours (seed=0, expansion=0.0) | SI2 (seed=0, expansion=0.12) | Ours (seed=1, expansion=0.0) | SI2 (seed=1, expansion=0.1) | Ours (seed=2, expansion=0.0) | SI2 (seed=2, expansion=0.15) | Ours (seed=3, expansion=0.0) | SI2 (seed=3, expansion=0.09) | Ours (seed=4, expansion=0.0) | SI2 (seed=4, expansion=0.14)