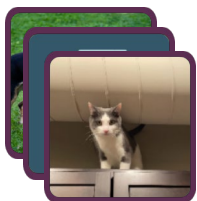
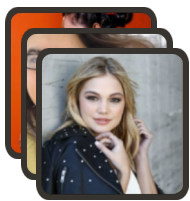


Auto

$N \times \text{image} + N \times \text{model} \rightarrow N \times \text{predictions}$



Animal Automatic
Matting



Human Automatic
Matting



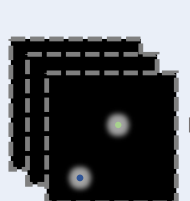
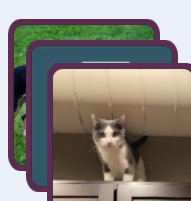
Natural Automatic
Matting

Aux-based

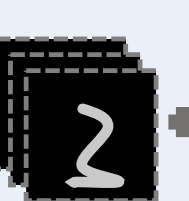
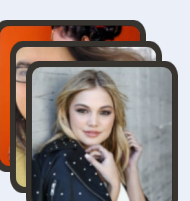
$N \times \text{image} + N \times \text{prompt} \rightarrow N \times \text{predictions}$



+



+



+



Foreground-agnostic Auxiliary-input based Matting

In-Context

$N \times \text{image} + 1 \times \text{prompt} + 1 \times \text{model} \rightarrow N \times \text{predictions}$



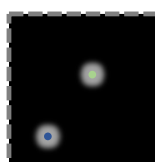
Reference



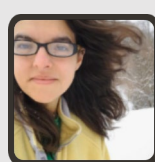
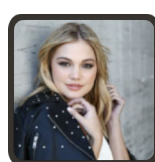
...



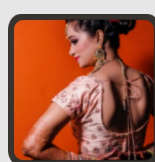
Target images



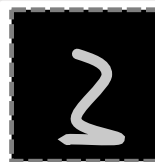
Reference



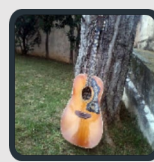
...



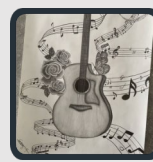
Target images



Reference



...

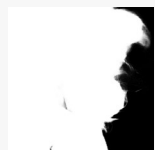


Target images

In-Context Matting



...



...



...

