

## Original Datasets

WMCA

PADISI

SIW

## Uni-class Datasets

Bonafide

Fakehead

Print

Glasses



Replay

Paper mask

Flexible mask

Rigid mask

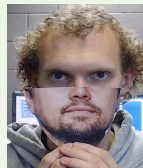


Partial Eye

Partial Mouth

Makeup

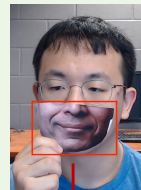
Tattoo



## Visual Gounding Datasets



label



[ 92, 339, 352, 507 ]

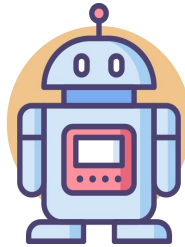
## System Prompt

### System Message:

"A chat between a curious user and an artificial intelligence assistant. The assistant gives helpful, detailed, and polite answers to the user's questions.

**USER:** <image>\n{prompt}

**ASSISTANT:** "..."



MLLM

## Dataset Construction Prompt

This is a **{real/spoof}** face picture **{simple description.(eg.The spoofing glasses are used to simulate the real human eye part and to deceive the face detection system.)}**.

Combined with the picture and the picture information I have provided, I hope you can help me generate a question and answer data. Here's what I hope:\n

Question: The contents are various one-sentence instructions about **{Question Type. (eg.whether the face is real or spoof.)}**.\n

Answer: The result of judgment in one paragraph. Give a definite answer about **{Question Type}**. \n

...  
Here is some examples of implementing the above requirements:\n**{examples}**:...

## MLLM Reponse

### Pretrain

Q: "<image>\n**What does the image depict?**"

A: The image is an artistic representation of a person's face, where their head has been replaced with a rigid mask made up of plastic material. ....

### Coarse-grained Classification

Q: "<image>\n**Can the image be considered a genuine portrait or a spoof face?**"

A: "The image is a spoof face,it shows a person is wearing a rigid mask that covers their entire face, likely made of plastic material, to create a spoofed effect."

### Fine-grained Classification

fineQ: "<image>\n**What is the type of facial attack shown in this image?**"

A: "The face spoof picture represents a Rigid Mask Attack. This is indicated by the presence of a rigid, plastic mask covering the face, creating a face spoofed effect."

### Reasoning

Q: "<image>\n**Why the face real or spoof?**"

A: ""According to the picture, the **"Face Facial Lighting"** is intense and creates an artificial glare...  
The **"Global Shape Consistency of Facial Features"** is ...  
The **"Sense of Depth and Three-Dimensionality"** is ...  
Finally, the **"Presence of Phone Screen or Paper Edges"** suggests that ..  
Analyzing these attributes, we can conclude that the face in the picture is a spoof."

### Localization

Q: "<image>\n**Can you locate the spoof area of the face in the image?**"

A: "The spoofing region is precisely identified at **[92.0, 339.0, 352.0, 507.0]** . Here, a mask is placed to mimic human facial attributes...."

## Pretrain Dataset Filter

CLIP

Similarity>15%

Fas-Pre 10k

Fas-SFT 45k

## SFT Dataset Augmentation



LLM

Here is a question and answer pair in a JSON format:

Question: {question}

Answer: {answer}

Suppose you are an expert in data augmentation. Please generate two new question and answer pairs in the same JSON format.

## SFT Dataset Filter



LLM

Whether the question and answer correspond to each other?



Special words ...

"as an AI language model, I cannot..."  
"It is difficult to determine..."  
"it is not possible to..."



Human

Human review and revision