

## A ICONBANK

The construction of IconBank involves the following steps:

- Crawl icon-label pairs from Flaticon.
- Leverage the state-of-the-art vision-language model (LLaVA [27]) to generate detailed descriptions for these icons (see Figure 10).
- Leverage CLIP [39] to filter low-quality samples whose icon and description similarity is lower than a given threshold.



Figure 10: Some samples of IconBank.

## B ICONBENCH

We leverage ChatGPT to generate 100 test concepts with this prompt: "I am designing an icon generation model. Please help me generate a total of 100 concepts which are used for model evaluation. 50 of



Figure 11: IconDM can create novel icons from depth images.

ID	Abstract Concepts	ID	Concrete Concepts
1	security	51	apple
2	globalization	52	house
3	inspiration	53	cat
4	health	54	dog
5	environment	55	tree
6	progress	56	car
7	freedom	57	flower
8	collaboration	58	star
9	growth	59	bicycle
10	efficiency	60	computer
11	strength	61	key
12	technology	62	ball
13	poison	63	rocket
14	privacy	64	map
15	entertainment	65	fire
16	engineering	66	bird
17	trend	67	guitar
18	verification	68	present
19	transformation	69	cake
20	support	70	badge
21	fairness	71	microphone
22	information	72	battery
23	system	73	bell
24	hope	74	train
25	goal	75	glasses
26	friendship	76	scissors
27	creativity	77	Mount Fuji
28	education	78	Eiffel Tower
29	wisdom	79	pyramid
30	sustainability	80	Sydney Opera House
31	leadership	81	sandwich
32	courage	82	playground
33	justice	83	park
34	celebration	84	bag
35	challenge	85	farm
36	opportunity	86	tent
37	speed	87	wallet
38	fusion	88	trophy
39	strategy	89	mortarboard
40	solution	90	beverage
41	risk	91	coffee
42	success	92	orange
43	failure	93	hamburger
44	curiosity	94	candy
45	respect	95	ice cream
46	impact	96	egg
47	confidence	97	fish
48	service	98	school
49	music	99	hospital
50	update	100	shoes

Table 4: The complete list of concepts in IconBench.

1277 *them are concrete concepts, such as apple, house, cat. The other 50*  
 1278 *are abstract concepts, such as security, globalization, etc. Ensure that*  
 1279 *these concepts have large coverage so that the evaluation better reflects*  
 1280 *model performance."* We observe that there are duplicate concepts  
 1281 in the output of ChatGPT, so we perform 2-3 rounds of this process  
 1282 to remove duplicate concepts and get the final list (see Table 4).

## 1335 C MORE RESULTS WITH CONTROLNET

1336 In addition to canny images, with the support of ControlNet [59],  
 1337 IconDM can also generate icons based on depth images (see Fig-  
 1338 ure 11). 1339

## 1340 D MORE QUALITATIVE RESULTS

1341 More qualitative results and comparison are shown in Figure 12  
 1342 and Figure 13, respectively. 1343



Figure 12: More qualitative results of IconDM.

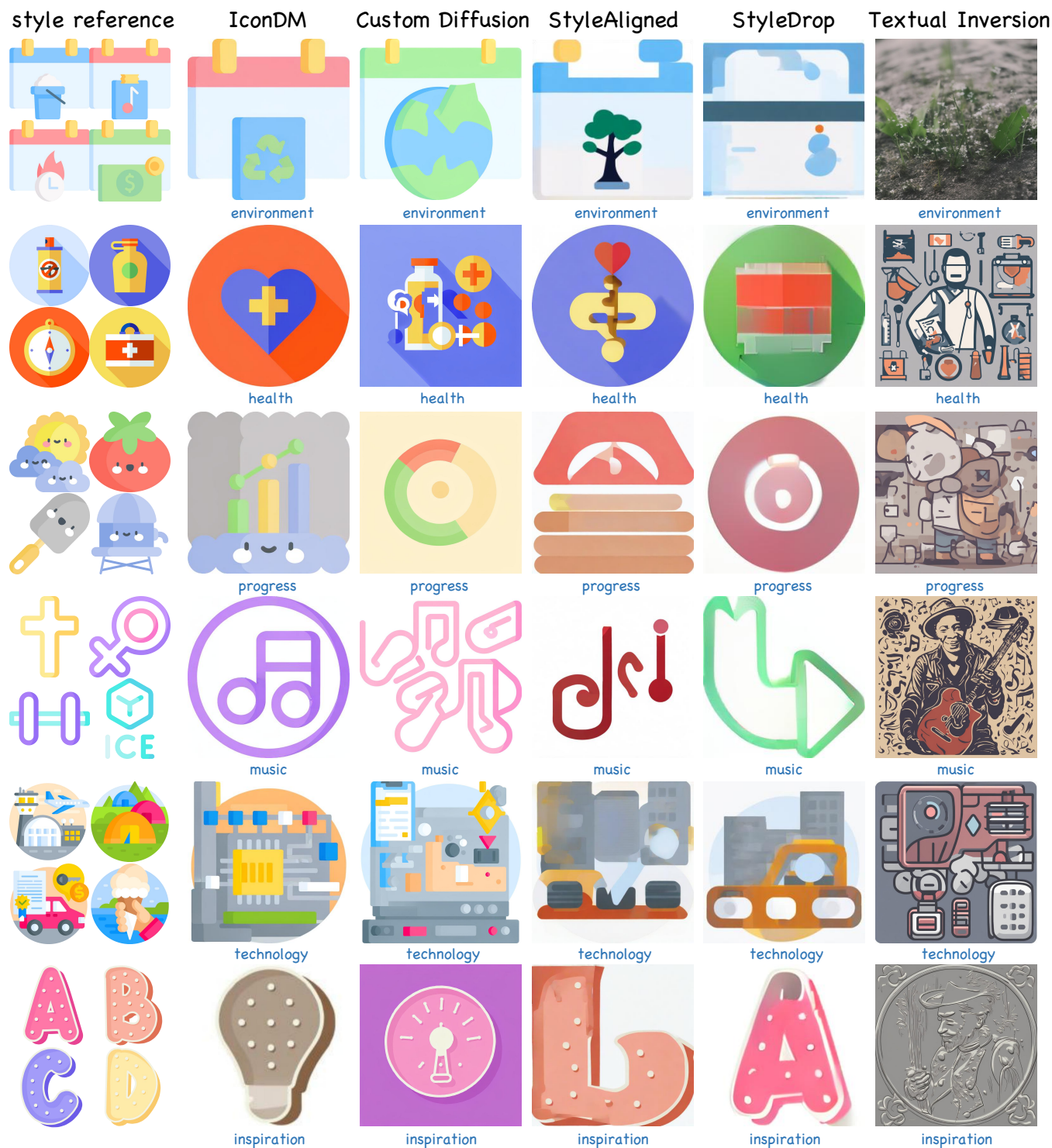


Figure 13: More qualitative comparison between IconDM and baselines.