

# Supplementary Materials: Self-Supervised Visual Preference Alignment

Anonymous Authors

## A RELATIONS WITH CONTRASTIVE LEARNING

As discussed before in Sec. 3.3 of the main paper, the optimized function in SeVa has strong relations with visual contrastive learning. In this section, we will derive a more general form of DPO loss that can be easily adapted to our SeVa pipeline. We first rewrite the InfoNCE loss here for clarity:

$$\mathcal{L}_{\text{in}} = -\log \frac{\exp(q \cdot k_+ / \tau)}{\exp(q \cdot k_+ / \tau) + \sum_i^n \exp(q \cdot k_-^i / \tau)}, \quad (14)$$

Inspired by Eq. 14 that multiple negative terms are involved, we can reformulate the preference distribution in Eq. 3 (in the main paper) as:

$$p_{\text{multi}}^* = p^*(y_c > Y_r | x) = \frac{\exp(r^*(x, y_c))}{\exp(r^*(x, y_c)) + \sum_{y_r \in Y_r} \exp(r^*(x, y_r))}, \quad (15)$$

where one positive preference data is paired with multiple rejected samples that are represented by a union  $Y_r$ . In this case, the preference database changes to:

$$\mathcal{D}^{\text{multi}} = \{x^{(j)}, y_c^{(j)}, Y_r^{(j)}\}_{j=1}^{N_d}. \quad (17)$$

Note that in SeVa pipeline, this new preference database in Eq. 17 can be easily constructed (e.g., we can do data augmentation sampling across range of available ones to obtain more negatives for  $Y_c$ ). Now, by considering the closed form in of the reward function in Eq. 6 of the main paper and utilizing samples in Eq. 17 (to do maximum likelihood estimation), we derive a more general form of DPO as:

$$\mathcal{L}_{\text{dpo}}^{\text{multi}} = -\mathbb{E}_{\mathcal{D}^{\text{multi}}} \left[ \log \sigma \left( \beta \log \frac{\pi_{\theta'}(y_c | x)}{\pi_{\text{ref}}(y_c | x)} - \sum_{y_r \in Y_r} \beta \log \frac{\pi_{\theta'}(y_r | x)}{\pi_{\text{ref}}(y_r | x)} \right) \right], \quad (18)$$

and optimize it using pseudo-constructed database by our SeVa pipeline. At present, this generalized form is out of the scope of this paper. We will leave this as future work to explore more potentials of our SeVa.

## B DATA CONSTRUCTION DETAILS

In Sec. 4.1 of the main paper, we have discussed the data construction process. Here, we want to emphasize the *counting* of our data instance. Note that for clarity in all our experiment, we choose the num of the unfiltered preference sample as the DPO instances we used. But please note that the actual preference data sent to DPO are much less than that. One is the SeVa’s filtering mechanism as discussed before (the filtering process in main paper’s Alg. 1 almost cuts down the instance num by half). The other is the data counting difference between LLaVA and SeVa: an instance of multi-turn

dialogue will contribute to about 2 raw image-question pairs in our SeVa. As such, the data used in SeVa *can be* far less counted if we ‘regard’ this 2 image-question pairs as only one data instance that is sourced from a multi-dialogue.

## C MORE VISUALIZATIONS

### C.1 Data augmentations samples

In Fig. 9, we visualize the data augmentation samples to show how the negatives responses are obtained. We choose original image, random-flip (‘RandFlip’), random-resized-crop (‘RRCrop’), moco augmentations and diffusion noise for better visualizations. We highlight those incorrect or incomplete answers with red color.

### C.2 GPT-4 prompt template

Please refer to Fig. 8 for the GPT-4 prompt template we adopted to evaluate the consistency score in Table 5 of the main paper.

### C.3 More examples of SeVa

Please refer to Fig. 10-11. In Fig. 10, we showcase more examples of SeVa-13B to illustrate its multi-modal comprehension ability, including complex reasoning about a wedding, chain-of-thought power to make a drink and the capability to read scientific figures. In Fig. 11, we compare SeVa-13B with more VLMs, namely LLaVA-1.5-13B and Qwen-VL-chat. We found our model demonstrate a detailed and faithful answers, which is on par with (or even better than) Qwen-VL-chat that possibly requires more training data source and larger model capacity.

## REFERENCES

- [1] Anand Mishra, Shashank Shekhar, Ajeet Kumar Singh, and Anirban Chakraborty. 2019. Ocr-vqa: Visual question answering by reading text in images. In *2019 international conference on document analysis and recognition (ICDAR)*. IEEE, 947–952.
- [2] Amanpreet Singh, Vivek Natarajan, Meet Shah, Yu Jiang, Xinlei Chen, Dhruv Batra, Devi Parikh, and Marcus Rohrbach. 2019. Towards vqa models that can read. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8317–8326.



## ChatGPT for Consistency Evaluation

You are an AI assistant to help me identify the consistency of two set of answers generated by 2 Vision-Language models with different weights. The answers are generated according to the input image (invisible in your system), and an asked question (the 'prompt' dict in each question answer pair). The input to you are shown in the following format:

[task-specific input to the GPT]

### # Illustrations for input-specific information

The "prompt" value in the dict are the same questions asked to the two models (as mentioned before), while the 'text' value in each dict are the answers to Model1 and Model2, respectively. Now please help me evaluate two metrics, called "Question consistency of model1", "Question consistency of model2" and "consistency between model1 and model2" with an integer score from 1 to 10 (1 is the least consistency, and the 10 is the most consistent). Note that DO NOT writing any code to evaluate, but only evaluate your own, that is in your own system!

### # Illustrations for output-specific information

Here is the meaning of the three metrics

"A1, Question consistency of model1": the answered information ("text") of the model1 with regard to the asked questions ("prompt").

"A2, Question consistency of model2": the answered information ("text") of the model2 with regard to the asked questions ("prompt").

"A3, consistency between model1 and model2": The consistency of model2's answer compared with model1's answer. If the model2 outputs very different answers from model1, you should give this metric a relatively low score, and vice versa.

You should evaluate the consistency between the following 3 aspects:

1. wording: is the words used in the sentence consistency of not (score from 1-10, with 1 the least)
2. sentiment: is the sentiment conveyed consistent? (score from 1-10, with 1 the least)
3. content: is the overall content of facts consistent? (score from 1-10, with 1 the least)

Now I will show you one paired input, please output the metrics as follows

### # Requirements for output format

You need to figure out the [score] in the [OUTPUT] format. Also, print the SIMPLIFIED FORMAT as follows, and fill the correct number in the corresponding metrics.

### # Output examples

OUTPUT: [evaluation-specific output format]

SIMPLIFIED FORMAT:

A1,[wording],[content],[sentiment],[overall]

A2,[wording],[content],[sentiment],[overall]

A3,[wording],[content],[sentiment],[overall]

The INPUT Prompt is:

Prompt [Show your input to GPT for evaluation of consistency score]

Figure 8: Prompt template to evaluate the consistency score. The numerical results are shown in Table 4 of the main paper.

233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290

Original Image

Q1: Who is the author of this book?  
Q2: What type of book is this?

A1: Matthew Houde  
A2: This is a travel book, specifically a guidebook, that provides information and recommendations for travelers visiting Costa Rica.

RandFlip

Q1: Who is the author of this book?  
Q2: What type of book is this?

A1: **Hans-Martin Muller**  
A2: This is a travel book, specifically a guidebook, that provides information and recommendation for travelers visiting Costa Rica.

RRCrop

Q1: Who is the author of this book?  
Q2: What type of book is this?

A1: **Linda K. Watts**  
A2: This is a travel book.

MOCO-Aug

Q1: Who is the author of this book?  
Q2: What type of book is this?

A1: Matthew Houde  
A2: This is a **a book** about Costa Rica.

Diffusion-Noise

Q1: Who is the author of this book?  
Q2: What type of book is this?

A1: Matthew **House**  
A2: This is a travel book, specifically focused on Costa Rica.

(a) data augmentations of OCRVQA [1] images

Original Image

Q1: Provide a one-sentence caption for the provided image.\nReference OCR token: 10

A1: A laptop and monitor on a desk with a picture of a soccer player with the number 10 on his shirt.</s>

RandFlip

Q1: Provide a one-sentence caption for the provided image.\nReference OCR token: 10

A1: A computer monitor with a picture of a soccer player on it and **the number 10 on the screen.**

RRCrop

Q1: Provide a one-sentence caption for the provided image.\nReference OCR token: 10

A1: A laptop and monitor with a **soccer game on the screen.**

MOCO-Aug

Q1: Provide a one-sentence caption for the provided image.\nReference OCR token: 10

A1: A laptop is sitting on a table next to a monitor that has a picture of a **soccer player on it.**

Diffusion-Noise

Q1: Provide a one-sentence caption for the provided image.\nReference OCR token: 10

A1: A laptop and monitor on a desk **with a 10 on the monitor.**

(b) data augmentations of TextVQA [2] images

291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348

Figure 9: Visualizations of the distorted images and their model answers. In-complete or in-correct tokens are highlighted with red colors.

349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406

407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464



Q: Why are the man and woman in white so close to each other? Answer it and give the rationale.



**SeVa-13B:**  
The man and woman in white are close to each other because they are a bride and groom celebrating their wedding. In the image, the man is holding a glass of champagne and toasting, while the woman is also holding a glass of champagne and drinking. This scene represents a joyful moment during their wedding reception, where the couple is sharing a toast and enjoying each other's company. The close proximity of the couple symbolizes their unity, love, and commitment to each other as they embark on their new journey as husband and wife.



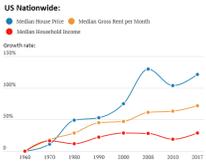
Q: How to make this?



**SeVa-13B:**  
To make this refreshing citrus drink with grapefruit balls and a sprig of mint, follow these steps:

- Gather ingredients:
  - Fresh grapefruit: peel and segment the grapefruit to remove the bitter pith and membranes. You can use a grapefruit zester or a sharp knife to do this.
  - Fresh orange: peel the orange to remove the outer layer and segment it to remove the seeds and pith.
  - Lime: peel the lime and remove seeds.
  - Mint leaves: fresh or dried, for garnish
  - Simple syrup (optional): you can make your own or use store-bought
  - Ice
  - Glass: a clear glass or a glass with a stem, preferably with a base wide enough to hold the grapefruit balls
  - Toothpick or skewer
- Prepare the grapefruit balls:
  - Combine the grapefruit segments, simple syrup (if using), and a squeeze of lime juice in a bowl.
  - Mix well and refrigerate for at least 30 minutes to allow the flavors to meld.
  - Form the grapefruit segments into balls using your hands or a small cookie scoop.
- Assemble the drink:
  - Fill the glass with ice.
  - Add a few grapefruit balls to the glass.
  - Garnish the drink with a sprig of mint and a toothpick or skewer to hold it in place.
  - Pour the prepared grapefruit juice over the ice and grapefruit balls.
  - Squeeze a wedge of lime over the drink and drop it in to release the juice.
  - Stir gently and serve immediately.

This citrus drink is perfect for warm weather or as a refreshing treat during any time of the year. Enjoy!



Q: Which year has the highest growth rate of median house price?



**SeVa-13B:**  
The year with the highest growth rate of median house price is 2000, as indicated by the blue line on the chart. This suggests that during that time, there was a significant increase in the median house price, which could be attributed to various factors such as economic growth, low-interest rates, and increased demand for housing.

Figure 10: Visualizations of complex visual question answering with our SeVa-13B models.

465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522



Q: What are the meals that I can cook with these?

**LLaVA-1.5-13B:** With the ingredients available in the refrigerator, you can prepare a variety of meals. The refrigerator contains several carrots, which can be used as a healthy snack or incorporated into dishes like salads, stir-fries, or soups. The milk can be used to make smoothies, pancakes, or creamy sauces for pasta dishes. The yogurt can be used as a base for dips, dressings, or as a topping for oatmeal or granola. The strawberries can be eaten on their own or used as a topping for desserts or fruit salads. The eggs can be used to make omelettes, frittatas, or scrambled eggs. The juice can be consumed as a refreshing beverage or used as an ingredient in recipes like marinades or glazes. The bread can be used to make sandwiches, toast, or as a base for various savory or sweet dishes. The butter can be used for cooking or baking, and the cream can be used in desserts or sauces. The strawberries and cream can be combined to make a delicious fruit salad or dessert. Overall, the refrigerator offers a diverse range of ingredients that can be used to create a variety of healthy and delicious meals.



**Qwen-VL-Chat:**

Based on the ingredients visible in the image, you could prepare a variety of meals. Here are some ideas:

- Yogurt Parfait: Layer Greek yogurt, fresh berries (strawberries and blueberries), and granola or honey for a healthy breakfast or snack.
  - Smoothie Bowl: Blend frozen mixed berries, banana, and almond milk to make a smoothie base. Top it with fresh fruit, yogurt, and granola for a nutritious breakfast or snack.
  - Carrot Salad: Mix shredded carrots with raisins, chopped apples, and a simple dressing made from olive oil, lemon juice, and honey. Serve as a side dish or light lunch.
  - Fruit Salad: Combine strawberries, blueberries, and other fruits like bananas, grapes, or pineapple for a refreshing dessert or snack.
  - Yogurt Dip: Mix Greek yogurt with garlic, lemon juice, and herbs like dill or parsley. Serve with carrot sticks or other vegetables for a healthy appetizer or snack.
- Remember to store your ingredients properly to maintain their freshness and quality.



**SeVa-13B:**

With the variety of food items in the refrigerator, you can prepare a range of meals and snacks. Some ideas include:

1. Fruit salad: Combine the strawberries and blueberries to create a refreshing fruit salad. You can also add other fruits like apples, oranges, or grapes for added variety.
2. Yogurt parfait: Use the yogurt as a base for a delicious yogurt parfait. Layer the yogurt with granola, fresh fruit, and nuts for a nutritious and satisfying breakfast or snack.
3. Carrot sticks with dip: Slice the carrots into sticks and serve them with a healthy dip, such as hummus, guacamole, or a yogurt-based dip.
4. Smoothie: Blend the yogurt, fruit, and any additional ingredients like nuts, seeds, or protein powder to create a nutritious and delicious smoothie.
5. Milk-based dessert: Use the milk to make a creamy dessert, such as a milkshake, pudding, or custard. You can also mix it with fruit or yogurt to create a refreshing dessert.
6. Oatmeal: Use the milk as a base for oatmeal, and add fruits, nuts, and other toppings for a hearty and filling breakfast.

These are just a few examples of the meals and snacks you can create using the food items in the refrigerator. With a little creativity and experimentation, you can come up with many more delicious and nutritious dishes.



Figure 11: Comparison of our SeVa-13B with other VLMs (LLaVA-1.5, Qwen-VL-Chat).

523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580