

## Supplementary Material of “Nested-Wasserstein Self-Imitation Learning for Sequence Generation”

### A More Details about WSIL

**Direct nested-Wasserstein Self-Imitation Learning** Direct Wasserstein self-imitation learning (WSIL-D) weights the original rewards with outputs from the behavior policy for sequences in the replay buffer  $\mathcal{B}$ . The sequences from the replay buffer are directly used as pseudo-samples to update the generator (Liang et al., 2018). Similarly, define  $r_{ns}(Y^s, \mathbf{Y}) \triangleq \sum_j T_j^s r_s(Y^s, Y_j)$ , with  $\mathbf{T}' = \{T_j^s\}$  the optimal weights.  $r$  to be the nested-Wasserstein reward between the sequence  $Y^s$  and ground-truth references  $\mathbf{Y}$ . The general objective (6) is then extended to be the objective for WSIL-D, as

$$J_D(\pi_\theta) \triangleq \mathbb{E}_{Y^s \sim \pi_{\theta, X}} [r(Y^s)] \quad (10)$$

$$+ \lambda \mathbb{E}_{Y^b \sim \pi_{\mathcal{B}, X}} [r_{ns}(Y^b, \mathbf{Y}) \pi_\theta(Y^b)], \quad (11)$$

where  $r$  is the original RL reward;  $r_{ns}$  is the nested-Wasserstein reward. Based on the objective of (11), we update the generator with standard RL loss and the self-imitation loss alternatively, with a hyperparameter  $\lambda$  that controls the update frequency:

$$\nabla_\theta J_D(\pi_\theta) \approx - \sum_{k=1}^K [(r(Y_k^s) - b) \nabla_\theta \log \pi_\theta(Y_k^s)]$$

$$- \lambda \sum_{j=1}^{K'} [(r_{ns}(Y_j^b, \mathbf{Y}) - b_s)_+ \nabla_\theta \log \pi_\theta(Y_j^b)] \quad (12)$$

where  $(\cdot)_+ \triangleq \max(\cdot, 0)$  and  $b_s$  and  $b$  are the baselines to reduce the variance of gradient estimates. In practice,  $(\cdot)_+$  means that WSIL-D only imitates the sequences in the replay buffer with the higher rewards. Intuitively, direct self-imitation implicitly imposes larger weights on good simulated data for training, to exploit good historical explorations. The main difference between WSIL-D and its indirect counterpart is that sequences from the replay buffer are not used to compute the self-imitation rewards, but used to evaluate the policy. Intuitively, WSIL-D changes the data distribution to explore the good history more efficiently.

### B Implementation Details

**Replay Buffer Construction** In our algorithm, a metric is required to be designed to select high-reward history demonstrations, which will be stored in the replay buffer  $\mathcal{D}$ . There are different ways for evaluating sentences:

---

#### Algorithm 2 IPOT for Wasserstein Rewards

---

- 1: **Input:** Feature vectors  $\boldsymbol{\mu} = \{\mathbf{z}_i\}_1^n$ ,  $\boldsymbol{\nu}' = \{\mathbf{z}'_j\}_1^m$  and generalized stepsize  $1/\lambda$ ,
- 2:  $\boldsymbol{\sigma} = \frac{1}{m} \mathbf{1}_m$ ,  $\mathbf{T}^{(1)} = \mathbf{1}_n \mathbf{1}_m^\top$
- 3:  $\mathbf{C}_{ij} = c(\mathbf{z}_i, \mathbf{z}'_j)$ ,  $\mathbf{A}_{ij} = e^{-\frac{\mathbf{C}_{ij}}{\lambda}}$
- 4: **for**  $t = 1, 2, 3, \dots$  **do**
- 5:  $\mathbf{Q} = \mathbf{A} \odot \mathbf{T}^{(t)}$  //  $\odot$  is Hadamard product
- 6: **for**  $k = 1, \dots, K$  **do**
- 7:  $\boldsymbol{\delta} = \frac{1}{n \mathbf{Q} \boldsymbol{\sigma}}$ ,  $\boldsymbol{\sigma} = \frac{1}{m \mathbf{Q}^\top \boldsymbol{\delta}}$
- 8: **end for**
- 9:  $\mathbf{T}^{(t+1)} = \text{diag}(\boldsymbol{\delta}) \mathbf{Q} \text{diag}(\boldsymbol{\sigma})$
- 10: **end for**
- 11: **Return**  $\langle \mathbf{T}, \mathbf{1} - \mathbf{C} \rangle$

---

*i)* For unconditional generation with synthetic data, following Chen et al. (2018), we adopt the negative log-likelihood (NLL) to measure model performance, as there exists an oracle data distribution. For this experiment, the replay buffer is constructed by generated sentences which achieved higher reward from the learned discriminator.

*ii)* For unconditional generation with real data, since we will use Test BLEU score and Self BLEU score for evaluating generated sentences, we maintain a single large replay buffer with BLEU-F1 score as the selection criteria to evaluate quality and diversity trade-off Gu et al. (2019). F1-BLEU score is defined as the geometry average of BLEU score and 1- Self-BLEU

$$\text{F1-BLEU} = \frac{2 \times \text{BLEU} \times (1 - \text{Self-BLEU})}{\text{BLEU} + (1 - \text{Self-BLEU})}. \quad (13)$$

*iii)* For conditional generation with captioning task, we maintain a small ( $K' = 5$  sequences) replay buffer for each conditional input; the replay buffer seems large, but we only need to store sequences of indexes, which is very efficient. Here we use the nested Wasserstein rewards as the metric.

*iv)* For conditional generation with non-parallel style transfer, we maintain a large replay buffer storing successfully transferred pairs, and we define a metric which considers both the accuracy and content preservation:  $p(\text{Right Style}) \times \text{BLEU}$ .

**Balance between RL and self-imitation** According to the theory of Wasserstein policy gradient Villani (2008),  $1/\lambda$  defined in Section (6) can be interpreted

as generalized decaying learning rate. With more explorations,  $\lambda$  becomes larger, and the algorithm should focus more on the self-imitated learning. In practice, we do one self-imitated learning update with every 10 RL training updates, and as training proceeds, we increase the frequency of self-imitation, and finally update the generator with one-step self-imitation followed with one-step standard RL training.

**The trick of soft-argmax** Recall that in sequence generation, one first samples a token based on the policy, then feeds its token embedding into the RNN to compute the logits of the next token, and repeat the above process based on the logits again until the stop token is generated. Instead of using the embedding of a sampled token, the soft-argmax trick feeds the RNN with the weighted average of the embeddings of most-likely tokens. In particular, let  $E$  be the word embedding matrix,  $g_t$  be the logits under the current policy and  $\mathbf{s}_t$  be the hidden state of the policy  $\pi_\theta$ . With the soft-argmax trick, the state vector is updated by

$$\tilde{y}_t = E \cdot \text{softmax}(g_t/\beta), \quad (14)$$

$$\tilde{\mathbf{s}}_t = h(\tilde{\mathbf{s}}_{t-1}, e(\tilde{y}_t)), \quad (15)$$

where  $0 < \beta < 1$  is the annealing factor, and in practice, we set  $\beta = 0.01$ .

**Discriminator implementation** In unconditional generation, instead of using policy gradient and the output of the discriminator as rewards, we use the soft-argmax trick Hu et al. (2017). Since the policy gradient is not stable enough and soft-argmax trick gives us better performance (See our extensive experiments).

**Nested-Wasserstein rewards implementation** In conditional generation, the Wasserstein rewards is implemented based on COCO test tools, and we use the fasttext Mikolov et al. (2018) as the fixed word embedding to compute the reward. In practice, we use  $K = 5$  with a hyper-parameter search from  $\{3, 5, 8, 10\}$ . *We will release this code, which is easy to use as other metrics.* For unconditional generation, we use the fixed learned word embedding via stop its gradient, where the embedding and the Wasserstein trust region are jointly optimized.

We conduct experiments on synthetic data similar to Yu et al. (2017), where our implementation is based on LeakGAN. The result is shown in Figure 3, where WSIL-I and WSIL-D show better performance than LeakGAN. Specifically, LeakGAN is not stable in the training and the Negative log-likelihood increases after 150 epochs. Compared with LeakGAN, WSIL-I and WSIL-D are more stable.

## C Experimental Setup

**Conditional text generation** We consider image captioning using the COCO dataset Lin et al. (2014), which contains 123,287 images in total, each of which is annotated with at least 5 captions. Following Karpathy’s split Karpathy and Fei-Fei (2015), 113,287 images are used for training and 5,000 images are used for validation and testing. We follow the implementation of the SCST approach (Rennie et al., 2016), and use extracted image tags (Gan et al., 2017; Wang et al., 2019) as image features (encoder). The learning rate of the generator is 0.0002, the maximum length of sequence is set to 25. For video captioning, the learning rate of the generator is 0.0001, the maximum length of sequence is set to 30. We use fixed image features and do not finetune the image encoder following previous work. A one-layer LSTM with 1024 units is used as the decoder. The word-embedding dimension is set to 512.

**Unconditional text generation** We use the COCO dataset Lin et al. (2014), in which most sentences are of length about 10. Since we consider unconditional text generation, only image captions are used as the training data. After preprocessing, the training dataset consists of 27,842 words and 417,126 sentences. We use 120,000 random sample sentences as the training set, and 10,000 as the test set. For the COCO dataset, the learning rate of the generator is 0.0002, the learning rate of the manager is 0.0002 (we follow the LeakGAN work), and the maximum length of sequence is set to 25.

Following Zhu et al. (2018), we use the News section in the EMNLP2017 WMT4 Dataset as our training data, which consists of 646,459 words and 397,726 sentences. After preprocessing, the training dataset contains 5,728 words and 278,686 sentences. The learning rate of the generator is 0.0002, the learning rate of the manager is 0.0002, and the maximum length of sequence is set to 50. The number of hidden units used in both the LSTM for the generator and the manager are set to 128. The dimension of the word embedding is 300. The discriminator is a CNN with its structure specified in Table 7.

**Settings of human evaluation** We perform human evaluation using Amazon Mechanical Turk, evaluating the text quality based on readability and meaningfulness (whether sentences make sense). We ask the worker to rate the input sentence with scores scaling from 1 to 5, with criterion listed in Table C. We require all the workers to be native English speakers, with approval rate higher than 95% and at least 100 assignments completed.

|  |
|--|
| Sequence to a scalar value                         |
| Input $300 \times$ Seq. Length Sequences           |
| (Kernel Size: Num( $\times 300$ ), Kernel Numbers) |
| (1, 100),(2, 200),(3, 200),(4, 200),(5, 200)       |
| (6, 100),(7, 100),(8, 100),(9, 100),(10, 100)      |
| (16, 160),(20, 160),(30, 160),(40,160)             |
| MLP output 1, ReLU                                 |

Table 7: Architecture of the discriminator.

| Scores    | Criterion  |
|-----------|--|
| 5 (Best)  | It is consistent, informative, grammatically correct.        |
| 4         | It is grammatically correct and makes sense.                 |
| 3         | It is mostly meaningful and with small grammatical error.    |
| 2         | It needs some time to understand and has grammatical errors. |
| 1 (Worst) | Meaningless, not readable.                                   |

Table 8: Human evaluation rating criterion.

| Dataset       | Train   | Test   | Vocabulary | Average Length |
|---------------|---------|--------|------------|----------------|
| Synthetic     | 10,000  | 10,000 | 5,000      | 20             |
| COCO captions | 120,000 | 10,000 | 27,842     | 11             |
| WMT News      | 278,686 | 10,000 | 5,728      | 28             |

Table 9: Brief description of the datasets used in unconditional text generation.

## D Generated Samples of Unconditional Text Generation

We show the generated samples of EMNLP NEWS2017 in Table 10, Table 11 and MS COCO in Table 12. Please Note all the samples are randomly selected from the generated sentences, without any human selection. It is obvious to see the diversity of LeakGAN is very poor in MSCOCO Captions, since it keeps generating sentences started with 'a'. Our proposed methods are more similar to the real data.

## E Generated Samples of Image Captioning

We show the generated samples of Image Captioning in Figure 7. We compares WSIL-D with SIL-D. We highlight benefits of using Wasserstein rewards, and put scores of each candidate.

| Methods   | Generated Examples   |
|-----------|--|
| Real Data | <p>But public opposition to the policy has been growing in other countries , and Austria on Wednesday announced an overall limit over the next four years of 130 , 000 - or the equivalent of 1 . 5 per cent of the population .</p> <p>This time , the government will put the draft to a referendum , which is expected in July though no date has been fixed .</p> <p>I feel that sometimes the people accept me the way I am and other times they don ' t accept me at all .</p> <p>For years the state told us we were crazy , that our water was safe , which wasn ' t true .</p> <p>It provides less accommodation of companies engaged in high - cost development and more reward for those that can lower their cost structures .</p> <p>When you win a title you gain confidence , and the supporters love you , because they want to win things as well .</p> <p>The combined value of the contracts is about \$ 8 . 3 million but could nearly double once additional funding is provided .</p> <p>He also imposed conditions on a release on bond that include being placed on an electronic monitor , drug testing and reporting weekly to authorities .</p> <p>Only then would a discussion begin within the Justice Department over whether to pursue any legal action against Clinton or anyone else involved in the matter .</p> <p>Andy Hall , an advocate for migrants who advised defence lawyers in the case , said the defence requested additional DNA - related documents from the prosecution but they were not provided .</p> <p>Wales put tickets for its three home matches - versus Scotland , France and Italy - on general sale back in October , with Scotland tickets now completely sold out .</p>  |
| LeakGAN   | <p>It ' s not easy but I have to be with the fact that the problem is probably : ' t like me ," he said .</p> <p>" This is the lie that Ted ' s campaign is built on ," Rubio said of his fellow challenge as the EU to vote for prime minister .</p> <p>The new rules mean that international companies will have to tell the country they operate in what they make in up companies do just over their year as they ' s having sex .</p> <p>The court said that the UK ' s biggest country could have an impact on the site , a very long time has been the only to three years left .</p> <p>There was one male friend , however , who admitted that to were in front of five minutes away from his home on the other side .</p> <p>The 32 - year - old reality star gave birth to their lawyers told the United States were the Republican - year - old girl who has a very high out .</p> <p>The team of Ohio State researchers set out to determine what they had " been " more " head of an " a " country or seven , according to the public of the incident .</p> <p>As a result , most people believed they were voting for his voice is only going to get the data right when he is there .</p> <p>The main thing for us is to keep it as a long - term - wide range of travel , that their calls for students or twice - and - she said .</p> <p>The committee also said some people decide to move as many as the highest - child coalition can get the little of better - and have done in the attacks , at the point when they are having the best chance of the victory .</p> <p>The report , however , was a child ' s first child in the ISIS commander , the second half since the past seven years , it has been No on the family who do not have a gun control .</p>   |
| WSIL-D    | <p>A report from Kings College London last year revealed that members of the UK armed forces are twice as likely to develop depression or anxiety than members of the general working population .</p> <p>We need to identify with him on a human level , to understand whatever he does in his job in Afghanistan he ' s also affected by stuff that happens at home .</p> <p>She said she was in the car park when Campbell climbed into the drivers ' seat of a vehicle , prompting her to offer him £ 20 to get a taxi instead .</p> <p>A report published by NHS England found it had failed to investigate hundreds of deaths over four months before the 2020 election and they did not want to even more common if they are .</p> <p>You need to be absolutely totally clear about which customers you are going to see a lot of people out .</p> <p>In the UK , parents , local authorities , charities , the media and politicians have all bought into the schools - can - fix - it narrative .</p> <p>The annual report , on behalf of the Welsh government , also found more people than ever are being treated .</p> <p>In his view , although he can be seen with a £ three million to expand its annual million to income out 4 per cent .</p> <p>It was the first day I fell in with the first year I ' ve ever been playing for a long - term plan for 45 per cent .</p> <p>And then I ran into him out a few months later and we started hanging out and now we are in a relationship with that we all .</p> <p>A decision from the ACT on the dispute between the national energy regulator and the power networks was due by December 22 , but the ACT advised before Christmas it could be up to three months late .</p> <p>" It certainly gave us a boost , it was like a late Christmas present but it was about it ," he said .</p>   |
| WSIL-I    | <p>He said he was using his executive powers as president because the US Congress has failed to address the problem .</p> <p>When I would make my meals for my family , I would double it and bring a meal of the year ' s heart ," she says , at the time .</p> <p>We accept all the recommendations for the Ministry of Justice in this report and are already taking action to implement them .</p> <p>This has been a dream scared , but for the long - term goal would stay be from class - to - the - quarter down .</p> <p>The Trump campaign will air the ad in the early - voting states of Iowa , New Hampshire , and South Carolina .</p> <p>Both winners said the crowds at this year ' s event seemed similar to last year , although official numbers found the four - day crowd was slightly smaller at just over 100 , 000 .</p> <p>But the one good thing we can take from this is it ' s happened quite early in the wet season and , what more people are .</p> <p>It ' s nice to know that I am wanted . I have lost a lot of confidence in myself over the last two days ," he said .</p> <p>The president responded that those criminals illegally purchase weapons from others who should ' ve been subject to background checks .</p> <p>I ' ve got worse since this started , I ' ve isolated myself even over the last couple of months .</p> <p>According to Swedish Radio , police want up to 2 , 500 more officers and 1 , 600 new civilian workers by the year 2020 .</p> <p>I don ' t know what the truth is and I don ' t , as a regular citizen , know how to find that information out .</p> <p>We might think we know where we ' re going , but the way ahead , and the path behind , when the show was to work out .</p> <p>He was told that he didn ' t even think he could have had information but to the evidence to make a couple of weeks .</p> <p>It ' s great that we hold ourselves back and we know about every January we had the best of the season we ' ll have just as to be the best in the world .</p> |

Table 10: Generated examples on EMNLP2017 WMT.

## Nested-Wasserstein Self-Imitation Learning for Sequence Generation

| Methods | Generated Examples   |
|---------|--|
| SeqGAN  | <p>Following the few other research and asked for " based on the store to protect older , nor this .<br/>           But there , nor believe that it has reached a the person to know what never - he needed .<br/>           The trump administration later felt the alarm was a their doctors are given .<br/>           We have been the time of single things what people do not need to get careful with too hurt after wells then .<br/>           If he was waited same out the group of fewer friends a more injured work under it .<br/>           It will access like the going on an " go back there and believe .<br/>           Premier as well as color looking to put back on a his is .<br/>           So , even though : " don ' t want to understand it at an opportunity for our work .<br/>           I was shocked , nor don ' t know if mate , don ' t have survived ,<br/>           So one point like ten years old , but a sure , nor with myself more people substantial .<br/>           And if an way of shoes of crimes the processes need to run the billionaire .<br/>           Now that their people had trained and people the children live an actor , nor what trump had .<br/>           However , heavily she been told at about four during an innocent person .</p>   |
| MLE     | <p>Two separate officials are making a statement for comment , and people believe that the technology had started the act with several thousand in a million new location .<br/>           It ' s just that this attack is not used to the water there ' s been a lot to gain in the middle of their water , she said on Monday .<br/>           It is the first time the media science shows that women are here to be married , but this will never be forgotten . I wouldn ' t have made it down for my money , but I ' m happy to stay on , he says .<br/>           They think more is really the most important place to do with that , because the educational situation will be on the way forward .<br/>           I had a long time and investigators have said that it will be the wrong decision to establish cases , he said .<br/>           We will be trying to work with both of us to vote for the , for the next cabinet to get to the bottom of the negotiations , he said .<br/>           He had no a proper question I thought I was going to host over such a long time , he added .<br/>           We ' ve become more competitive , because it ' s a great year and we ' re going to do that .<br/>           You therefore have to understand the way our response goes the light we will on on this , the source said .<br/>           It wasn ' t the first time I went a little closer than I ' d had so we would do as a minimum .<br/>           He ' s not always played strong football and that ' s why he ' s ready to reach better goals and improve .<br/>           I think it ' s quite a different role , but we ' re never going to come out as we ' ve been too older or so .</p>   |
| SIL-D   | <p>He had a couple of clear , he had to be able to lead to be after he was seen as a result from the kind of type of involvement of both .<br/>           The company has said the final phase in its special group will be the police , they have to believe it .<br/>           But instead of helping her 15 years , but the most of permission down from some of the Republicans , he said , and he wants to be an all one to the next in - one - a Republican debate .<br/>           Go out and share the main entry of all the Syrian civil war , should to confirm the difficult , make response to the wrong end of the financial crisis .<br/>           To have a good group of quality quality abuse is a route - and is still there to yourself health , and how to feel is going to put on the roads a day , it ' s fine .<br/>           At some point it is , he has come to work hard for a few minutes to get the right up and they ' re not going . He thought that at the time she had done to take a couple of hours again before she ' s emails .<br/>           The most recent trade numbers had 3 . 4 percent of one in the national sector in the first few months .<br/>           He said : ' It is entirely possible that there ' s some people who are going to get to earn it up ," he said .<br/>           I think my business is very good very talented , and you are , and that the things that you can ' t teach , you ' re going to be fair - you need to .<br/>           By the time you had tens of some parts of the public ; it ' s not going to happen in the next few months .<br/>           " I think the show that I am not doing it is going to be a big story that he can ' t hope it .</p>  |
| SIL-I   | <p>But if they ' re not willing to do that , as we are working with more things on how to use it that .<br/>           But the business , which is due to be a report of London ' s New York Times , which is putting a lower tax in growth .<br/>           The man was not wearing the offered that ' s afford to being taken to the Royal Victoria Hospital and a third of them in line in the end with the same - one ' s .<br/>           A second man , aged 57 , had not been due to work in the city for the UK and in the end of the party .<br/>           I ' m just as a leader in the starting game to be anything about the balance ," he says .<br/>           The company has admitted the problems stem from an effort launched in 2005 to gain an interest to being an light - by Facebook said .<br/>           Apple has come at their point against people in the next five months and the next state ' s able to get the best way that we ' re going to make those at home .<br/>           He was one of the key moments , no case after the U . S . would get the right to doing it but could have been much in the summer .<br/>           It ' s a great look at all , and this means the is high and risk at one time even before the French , that is in the world , a good interest , it ' s taken very long .<br/>           The U . S . Energy Information Administration forecasts that the average price of people would be allowed to the European Union .<br/>           The top - ranked Djokovic has now beaten Murray in four of children who was going to back Southern California and headed to Britain ' s address and that the best of the party was to the show .<br/>           If we could find a way of starting out of the season it is also good to the way and they are in good and the manager , is playing better than those who sell .</p> |

Table 11: Generated examples on EMNLP2017 WMT.

| Method    | Generated Examples   |
|-----------|--|
| Real Data | <p>a surfer a woman and a child walk on the beach .<br/> a few people sit on a dim transportation system .<br/> a person protected from the rain by their umbrella walks down the road .<br/> the bathroom with a toilet has an interesting sink .<br/> two women preparing food in a kitchen one at the sink and one at the table .<br/> a white kitchen in a home with the light on .<br/> a cat stuck in a car with a slightly opened window .<br/> two bicycles and a woman walking in front of a shop .<br/> green tiled backsplash highlighted by low overhead lighting .<br/> a bicycle is parked by a bench at night .<br/> a brown horse is grazing grass near a red house .</p>  |
| LeakGAN   | <p>a bike parked in a bunch of grass on a sidewalk in front of a yellow and a yellow bus on a road .<br/> a dog is jumping high in the air .<br/> the bathroom is clean and ready for us to use .<br/> a woman in a bikini rides a skateboard .<br/> a bathroom with a mirror and a picture on the wall above .<br/> a bathroom with a toilet and a shower .<br/> a cat sitting on the roof of an old car .<br/> a bathroom with a toilet and a bathtub .<br/> a couple of people walking across a street holding an umbrella .<br/> a man sitting in front of a laptop computer on a counter .</p>  |
| WSIL-D    | <p>a large bathroom with a long marble shower .<br/> a bath and sink in a room with a large mirror .<br/> there is a woman that is sitting in the sink while the photo of a dog .<br/> white glass table sitting on top of a living room .<br/> woman in a blue dress sitting on a city street talking on a telephone .<br/> a person is taking a flash photo in a mirror .<br/> a bathroom sink with a mirror just above it .<br/> two guys are talking in a field with a blue bike in front of it as a train car .<br/> a nice bathroom with a standalone shower and a shower curtain .<br/> the corner of a rest room with toilet paper .<br/> a boy holding some yellow umbrella next to a street .<br/> some tables in a small wooden kitchen area .</p>  |
| WSIL-I    | <p>cat standing in sink and another woman in black tiled floor .<br/> a bathroom with tiled walls has a mirror on the wall .<br/> a black and white cat in a bathroom sink .<br/> the man is standing on his bike with the beach behind him .<br/> a person riding a long board down a road in front of a parked car .<br/> a bicycle and some pictures on the street corner with the car .<br/> the bathroom tub with ceramic tub has a glass door .<br/> a large modern lighted space with bath tub .<br/> the kitchen is preparing an elaborate appliances it .<br/> a guy jumping high in the air with people in around around .<br/> this family a man talks on his cell phone .<br/> a public toilet with the seat up in a bathroom .<br/> this kitchen with white cupboards and stainless steel oven in someones home .</p> |

Table 12: Generated examples on COCO.







| Images  | References  | Hypothesis  |
|---|---|---|
|    | <ol style="list-style-type: none"> <li>1. youngster on a skateboard <b>trying simple tip</b> up stunt</li> <li>2. young skateboarder displaying skills on sidewalk near field</li> <li>3. a person on a skateboard on a street</li> <li>4. a young boy is <b>performing tricks</b> on a skateboard</li> <li>5. a little boy that is jumping a skateboard</li> </ol>                         | a man riding a skateboard on a ramp<br>CIDEr: 1.23757203331<br>Wasserstein: 64.5707599929                           |
|   |   | a man <b>doing a trick</b> on a skateboard on a ramp<br>CIDEr: 1.01380281793<br>Wasserstein: 65.7397472556          |
|    | <ol style="list-style-type: none"> <li>1. tennis player holding a tennis racket on a court</li> <li>2. two people playing a match of tennis on a court</li> <li>3. a man with a tennis racket walks from the net of a tennis court</li> <li>4. a man standing on a tennis court holding a racquet</li> <li>5. a man in shorts and a long sleeve shirt playing tennis</li> </ol>             | a man standing on a tennis court holding <b>a</b><br>CIDEr: 2.92170011743<br>Wasserstein : 73.3066824325            |
|   |   | a man standing on a tennis court holding <b>a racket</b><br>CIDEr: 2.75596938069<br>Wasserstein: 75.6341653678      |
|   | <ol style="list-style-type: none"> <li>1. a fire hydrant on a sidewalk near a street</li> <li>2. there is a fire hydrant on the side of the road</li> <li>3. a fire hydrant sitting on the side of a road</li> <li>4. a <b>green</b> fire hydrant siting by a yellow pole</li> <li>5. a fire hydrant that is on a sidewalk</li> </ol>   | a fire hydrant on the side <b>of a</b><br>CIDEr: 4.20074054971<br>Wasserstein: 75.7075221289                        |
|   |   | a <b>green</b> fire hydrant on the side of a street<br>CIDEr: 3.45013175762<br>Wasserstein: 76.4022286692           |
|  | <ol style="list-style-type: none"> <li>1. a full plate full of delicious food sets on top of the table</li> <li>2. plate with dinner items on it on a white table cloth</li> <li>3. a plate with <b>chicken scallops</b> pasta and other food items</li> <li>4. a plate of food and a fork on a table</li> <li>5. a white plate topped with <b>chicken coleslaw and potatoes</b></li> </ol> | a plate of food on a table <b>with an</b><br>CIDEr: 1.15756194919<br>Wasserstein: 66.0165233321                     |
|   |   | a plate of food with <b>meat and vegetables</b> on a table<br>CIDEr: 0.778356131748<br>Wasserstein: 67.3537985559   |
|  | <ol style="list-style-type: none"> <li>1. one man rides a skateboard down an empty pool while another man watches</li> <li>2. two men in a drained pool one on a skateboard the other holding a skateboard</li> <li>3. two guys are skateboarding in an empty pool</li> <li>4. some teenage boys skateboard int he park</li> <li>5. a person on a skateboard on a ramp</li> </ol>           | a man riding a skateboard <b>on a</b> skate park<br>CIDEr: 0.959694148787<br>Wasserstein: 61.8797805118             |
|   |   | a man riding a skateboard <b>in a</b> skate park<br>CIDEr: 0.558768974896<br>Wasserstein: 62.2425281912             |
|  | <ol style="list-style-type: none"> <li>1. a giraffe that is laying on the ground</li> <li>2. giraffe lays down in an enclosure</li> <li>3. some giraffes dirt trees fence and a building</li> <li>4. a giraffe sitting down while another giraffes eats on branches</li> <li>5. three giraffes one seated on some dirt the other two standing</li> </ol>                                    | a group of giraffes standing next to a giraffe<br>CIDEr: 1.0495772431<br>Wasserstein: 57.0770483353                 |
|   |   | a group of giraffes standing next to a giraffe <b>in a zoo</b><br>CIDEr: 0.75825020238<br>Wasserstein: 58.283065908 |

Figure 7: Image captioning examples on COCO. Top: SIL-D; Bottom: WSIL-D. The examples are shown to highlight the benefits given by the Wasserstein rewards. As a gram-based hard-matching metric, CIDEr rewards focus more on the locality fluency and may render incomplete sentences. Wasserstein rewards focus more on semantic matching. WSIL provides a natural way to combine both benefits.