

Baseline model highest Bleu -> 16.67

2-channel decoder hyperparameter optimizations:

| num_enc | num_dec | num_heads | batch_sz | weight_reco g | Highest BLEU |
|---------|---------|-----------|----------|------------------|-----------------|
| 3 | 3 | 8 | 64 | 0.7 | 17.35 |
| 2 | 3 | 8 | 32 | 0.1 | 18.77 |
| 3 | 2 | 4 | 32 | 0.7 | 18.763 |
| 3 | 3 | 4 | 32 | 0.7 | 19.31 |
| 2 | 3 | 4 | 48 | 0.7 | 18.56 |
| 3 | 3 | 4 | 48 | 0.1 | 17.97 |
| 2 | 3 | 4 | 64 | 0.1 | 17.07 |
| 4 | 2 | 8 | 64 | 1.0 | 17.44 |
| 2 | 3 | 8 | 48 | 0.4 | 18.82 |

| | | | | | |
|---|---|---|----|-----|-------|
| 3 | 3 | 4 | 48 | 0.4 | 18.36 |
| 3 | 3 | 4 | 64 | 0.7 | 18.19 |
| 3 | 3 | 4 | 48 | 1.0 | 18.33 |
| 3 | 3 | 8 | 64 | 1.0 | 17.97 |
| 3 | 2 | 8 | 48 | 0.7 | 18.4 |

| | | | | | |
|---|---|---|----|-----|-------|
| 2 | 3 | 8 | 32 | 0.4 | 18.67 |
| | | | | | |

From above examples it is clear that test bleu is good when:

1. num_dec>num_enc
2. Weight_of_recognition loss is between (0.4,0.7)
3. Batch_size <=32

-->To make the test bleu better, I used another channel after the shared decoder layers. At the final decoder layer there are 3 channels performing 3 different tasks. First one is for SLT, another one is for generation of the glosses and another one is for de-en translation. Firstly I translated captions of the train dataset of German sentences to English using a model submitted by fairseq in WMT-19. I trained the transformer model to jointly perform SLT, SLR and en-de translation by having 5 shared decoder layers and 3 task-specific layers on top. This model gives a BLEU of 20.8 without hyper-parameter optimization.

SOTA -> 21.32

FURTHER EXPS:

Decoder split with recognition objective (loss_weight=0.7)-->

L_r = 0.7

L_{r_ctc} = 0.0

L_{trans} = 0.0

BATCH_SIZE = 32, num_enc_dec = 3,3...beam_size = 1 to 10

| BLEU-1 | BLEU-2 | BLEU-3 | BLEU-4 | ROUGE | SOTA(bleu-4) |
|--------|--------|--------|--------|-------|--------------|
| 44.66 | 32.78 | 25.62 | 21.08 | 46.06 | 21.32 |

EXP_____1:

hyperparameters:

lr_ctc_recog = 5.0

lr_dec_split_recog = 0.7

lr_dec_split_trans_de_en = 0.2

num_layers_enc=num_layers_dec=3

dropout = 0.1

2020-12-24 15:36:49,884 Training ended since there were no improvements in the last learning rate step: 0.000082

2020-12-24 15:36:49,885 Best validation result at step 37600: 22.29 eval_metric.

2020-12-24 15:37:29,056 -----

2020-12-24 15:37:29,056 [DEV] partition [RECOGNITION] experiment [BW]: 1

2020-12-24 15:37:50,516 finished in 21.4599s

2020-12-24 15:37:50,517 *****

2020-12-24 15:37:50,517 [DEV] partition [RECOGNITION] results:

New Best CTC Decode Beam Size: 1

WER 29.38 (DEL: 12.01, INS: 5.47, SUB: 11.90)

2020-12-24 15:37:50,517 *****

2020-12-24 15:37:50,517 -----

2020-12-24 15:37:50,517 [DEV] partition [RECOGNITION] experiment [BW]: 2

2020-12-24 15:38:31,693 finished in 41.1760s

2020-12-24 15:38:31,694 *****

2020-12-24 15:38:31,694 [DEV] partition [RECOGNITION] results:

New Best CTC Decode Beam Size: 2

WER 29.28 (DEL: 11.69, INS: 5.58, SUB: 12.01)

2020-12-24 15:38:31,694 *****

2020-12-24 15:38:31,694 -----

2020-12-24 15:38:31,694 [DEV] partition [RECOGNITION] experiment [BW]: 3

2020-12-24 15:39:35,543 finished in 63.8486s

2020-12-24 15:39:35,543 *****

2020-12-24 15:39:35,543 [DEV] partition [RECOGNITION] results:

New Best CTC Decode Beam Size: 3

WER 29.09 (DEL: 11.64, INS: 5.55, SUB: 11.90)

2020-12-24 15:39:35,543 *****

2020-12-24 15:39:35,543 -----

2020-12-24 15:39:35,543 [DEV] partition [RECOGNITION] experiment [BW]: 4

2020-12-24 15:41:01,551 finished in 86.0076s

2020-12-24 15:41:01,552 *****

2020-12-24 15:41:01,552 [DEV] partition [RECOGNITION] results:

New Best CTC Decode Beam Size: 4

WER 29.04 (DEL: 11.58, INS: 5.58, SUB: 11.88)

2020-12-24 15:41:01,552 *****

2020-12-24 15:41:01,552 -----
2020-12-24 15:41:01,552 [DEV] partition [RECOGNITION] experiment [BW]: 5
2020-12-24 15:42:51,453 finished in 109.9015s
2020-12-24 15:42:51,454 -----
2020-12-24 15:42:51,454 [DEV] partition [RECOGNITION] experiment [BW]: 6
2020-12-24 15:45:04,221 finished in 132.7665s
2020-12-24 15:45:04,221 -----
2020-12-24 15:45:04,221 [DEV] partition [RECOGNITION] experiment [BW]: 7
2020-12-24 15:47:39,513 finished in 155.2915s
2020-12-24 15:47:39,514 *****
2020-12-24 15:47:39,514 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 7
 WER 29.01 (DEL: 11.56, INS: 5.55, SUB: 11.90)
2020-12-24 15:47:39,514 *****
2020-12-24 15:47:39,514 -----
2020-12-24 15:47:39,514 [DEV] partition [RECOGNITION] experiment [BW]: 8
2020-12-24 15:50:35,850 finished in 176.3363s
2020-12-24 15:50:35,851 -----
2020-12-24 15:50:35,851 [DEV] partition [RECOGNITION] experiment [BW]: 9
2020-12-24 15:53:54,967 finished in 199.1157s
2020-12-24 15:53:54,967 -----
2020-12-24 15:53:54,967 [DEV] partition [RECOGNITION] experiment [BW]: 10
2020-12-24 15:57:37,511 finished in 222.5432s
2020-12-24 15:57:37,511
=====

2020-12-24 15:57:58,396 [DEV] partition [Translation] results:
 New Best Translation Beam Size: 1 and Alpha: -1
 BLEU-4 22.29 (BLEU-1: 48.68, BLEU-2: 35.25, BLEU-3: 27.36,
BLEU-4: 22.29)
 CHRF 45.52 ROUGE 48.05
2020-12-24 15:57:58,397 -----

2020-12-24 16:00:24,427 [DEV] partition [Translation] results:
 New Best Translation Beam Size: 2 and Alpha: -1
 BLEU-4 22.62 (BLEU-1: 47.89, BLEU-2: 34.96, BLEU-3: 27.43,
BLEU-4: 22.62)
 CHRF 45.06 ROUGE 48.25
2020-12-24 16:00:24,427 -----

2020-12-24 16:01:07,560 [DEV] partition [Translation] results:
 New Best Translation Beam Size: 2 and Alpha: 1
 BLEU-4 22.69 (BLEU-1: 48.13, BLEU-2: 35.14, BLEU-3: 27.56,
BLEU-4: 22.69)
 CHRF 45.24 ROUGE 48.34
2020-12-24 16:01:07,561 -----

2020-12-24 16:01:28,971 [DEV] partition [Translation] results:

New Best Translation Beam Size: 2 and Alpha: 2
BLEU-4 22.77 (BLEU-1: 48.21, BLEU-2: 35.25, BLEU-3: 27.66,
BLEU-4: 22.77)

CHRF 45.28 ROUGE 48.39

2020-12-24 16:01:28,971 -----

2020-12-24 16:06:54,221 [DEV] partition [Translation] results:

New Best Translation Beam Size: 4 and Alpha: 2

BLEU-4 22.80 (BLEU-1: 47.64, BLEU-2: 35.20, BLEU-3: 27.65,
BLEU-4: 22.80)

CHRF 45.18 ROUGE 48.69

2020-12-24 16:06:54,222 -----

2020-12-24 16:09:53,972 [DEV] partition [Translation] results:

New Best Translation Beam Size: 5 and Alpha: 2

BLEU-4 22.86 (BLEU-1: 47.39, BLEU-2: 35.12, BLEU-3: 27.67,
BLEU-4: 22.86)

CHRF 45.02 ROUGE 48.69

2020-12-24 16:09:53,972 -----

2020-12-24 16:19:56,604 [DEV] partition [Translation] results:

New Best Translation Beam Size: 8 and Alpha: 2

BLEU-4 22.87 (BLEU-1: 47.32, BLEU-2: 35.18, BLEU-3: 27.75,
BLEU-4: 22.87)

CHRF 44.99 ROUGE 48.87

2020-12-24 16:19:56,605 -----

2020-12-24 16:23:40,240 [DEV] partition [Translation] results:

New Best Translation Beam Size: 9 and Alpha: 2

BLEU-4 22.90 (BLEU-1: 47.42, BLEU-2: 35.21, BLEU-3: 27.77,
BLEU-4: 22.90)

CHRF 45.19 ROUGE 48.96

2020-12-24 16:23:40,240 -----

2020-12-24 16:37:56,223 *****

2020-12-24 16:37:56,223 [DEV] partition [Recognition & Translation] results:

Best CTC Decode Beam Size: 7

Best Translation Beam Size: 9 and Alpha: 2

WER 29.01 (DEL: 11.56, INS: 5.55, SUB: 11.90)

BLEU-4 22.90 (BLEU-1: 47.42, BLEU-2: 35.21, BLEU-3: 27.77,
BLEU-4: 22.90)

CHRF 45.19 ROUGE 48.96

2020-12-24 16:37:56,223 *****

2020-12-24 16:41:04,555 [TEST] partition [Recognition & Translation] results:

Best CTC Decode Beam Size: 7

Best Translation Beam Size: 9 and Alpha: 2

WER 28.46 (DEL: 9.72, INS: 5.38, SUB: 13.36)

BLEU-4 22.05 (BLEU-1: 47.07, BLEU-2: 34.57, BLEU-3: 26.96,
BLEU-4: 22.05)

CHRF 45.24 ROUGE 48.25

2020-12-24 16:41:04,556 *****

EXP _____ 2:

hyperparameters:

lr_ctc_recog = 0.0

lr_dec_split_recog = 0.7

lr_dec_split_trans_de_en = 0.0

num_layers_enc=num_layers_dec=3

dropout = 0.1

2020-12-23 04:21:29,919 Training ended since there were no improvements in the last learning rate step: 0.000240

2020-12-23 04:21:29,921 Best validation result at step 25600: 20.00 eval_metric.

2020-12-23 04:22:35,317

=====

2020-12-23 04:22:43,457 [DEV] partition [Translation] results:

New Best Translation Beam Size: 1 and Alpha: -1

BLEU-4 20.00 (BLEU-1: 45.47, BLEU-2: 32.18, BLEU-3: 24.72,
BLEU-4: 20.00)

CHRF 43.11 ROUGE 45.58

2020-12-23 04:22:43,457 -----

2020-12-23 04:23:40,398 [DEV] partition [Translation] results:

New Best Translation Beam Size: 2 and Alpha: -1

BLEU-4 20.51 (BLEU-1: 45.13, BLEU-2: 32.56, BLEU-3: 25.22,
BLEU-4: 20.51)

CHRF 43.28 ROUGE 46.03

2020-12-23 04:23:40,398 -----

2020-12-23 04:23:58,698 [DEV] partition [Translation] results:

New Best Translation Beam Size: 2 and Alpha: 1

BLEU-4 20.81 (BLEU-1: 45.49, BLEU-2: 32.89, BLEU-3: 25.54,
BLEU-4: 20.81)

CHRF 43.55 ROUGE 46.30

2020-12-23 04:23:58,698 -----

2020-12-23 04:26:20,270 [DEV] partition [Translation] results:

New Best Translation Beam Size: 4 and Alpha: 1

BLEU-4 20.87 (BLEU-1: 45.17, BLEU-2: 32.82, BLEU-3: 25.55, BLEU-4: 20.87)

CHRF 43.38 ROUGE 47.03

2020-12-23 04:26:20,270 -----

2020-12-23 04:27:45,901 [DEV] partition [Translation] results:

New Best Translation Beam Size: 5 and Alpha: 1

BLEU-4 21.20 (BLEU-1: 45.14, BLEU-2: 33.02, BLEU-3: 25.86, BLEU-4: 21.20)

CHRF 43.50 ROUGE 47.35

2020-12-23 04:27:45,901 -----

2020-12-23 04:29:19,757 [DEV] partition [Translation] results:

New Best Translation Beam Size: 6 and Alpha: 1

BLEU-4 21.20 (BLEU-1: 45.20, BLEU-2: 33.12, BLEU-3: 25.90, BLEU-4: 21.20)

CHRF 43.46 ROUGE 47.41

2020-12-23 04:29:19,757 -----

2020-12-23 04:29:33,792 [DEV] partition [Translation] results:

New Best Translation Beam Size: 6 and Alpha: 2

BLEU-4 21.25 (BLEU-1: 45.44, BLEU-2: 33.29, BLEU-3: 25.99, BLEU-4: 21.25)

CHRF 43.37 ROUGE 47.11

2020-12-23 04:29:33,793 -----

2020-12-23 04:31:01,364 [DEV] partition [Translation] results:

New Best Translation Beam Size: 7 and Alpha: 1

BLEU-4 21.32 (BLEU-1: 45.01, BLEU-2: 33.09, BLEU-3: 25.95, BLEU-4: 21.32)

CHRF 43.43 ROUGE 47.34

2020-12-23 04:31:01,364 -----

2020-12-23 04:31:16,555 [DEV] partition [Translation] results:

New Best Translation Beam Size: 7 and Alpha: 2

BLEU-4 21.36 (BLEU-1: 45.35, BLEU-2: 33.31, BLEU-3: 26.08, BLEU-4: 21.36)

CHRF 43.33 ROUGE 47.12

2020-12-23 04:31:16,556 -----

2020-12-23 04:32:51,765 [DEV] partition [Translation] results:

New Best Translation Beam Size: 8 and Alpha: 1

BLEU-4 21.45 (BLEU-1: 44.97, BLEU-2: 33.09, BLEU-3: 26.02, BLEU-4: 21.45)

CHRF 43.47 ROUGE 47.31

2020-12-23 04:32:51,765 -----

2020-12-23 04:33:08,199 [DEV] partition [Translation] results:

New Best Translation Beam Size: 8 and Alpha: 2

BLEU-4 21.51 (BLEU-1: 45.28, BLEU-2: 33.33, BLEU-3: 26.16,
BLEU-4: 21.51)

CHRF 43.36 ROUGE 46.96

2020-12-23 04:33:08,200 -----

2020-12-23 04:37:04,201 [DEV] partition [Translation] results:

New Best Translation Beam Size: 10 and Alpha: 1

BLEU-4 21.61 (BLEU-1: 44.85, BLEU-2: 33.17, BLEU-3: 26.16,
BLEU-4: 21.61)

CHRF 43.20 ROUGE 47.28

2020-12-23 04:37:04,201 -----

2020-12-23 04:38:22,033 *****

2020-12-23 04:38:22,033 [DEV] partition [Recognition & Translation] results:

Best CTC Decode Beam Size: -1

Best Translation Beam Size: 10 and Alpha: 1

WER -1.00 (DEL: -1.00, INS: -1.00, SUB: -1.00)

BLEU-4 21.61 (BLEU-1: 44.85, BLEU-2: 33.17, BLEU-3: 26.16,
BLEU-4: 21.61)

CHRF 43.20 ROUGE 47.28

2020-12-23 04:38:22,034 *****

2020-12-23 04:38:43,937 [TEST] partition [Recognition & Translation] results:

Best CTC Decode Beam Size: -1

Best Translation Beam Size: 10 and Alpha: 1

WER -1.00 (DEL: -1.00, INS: -1.00, SUB: -1.00)

BLEU-4 21.08 (BLEU-1: 44.66, BLEU-2: 32.78, BLEU-3: 25.62,
BLEU-4: 21.08)

CHRF 43.80 ROUGE 46.06

2020-12-23 04:38:43,938 *****

EXP_____3:

hyperparameters:

lr_ctc_recog = 5.0

lr_dec_split_recog = 0.0

lr_dec_split_trans_de_en = 0.0

num_layers_enc=num_layers_dec=3

dropout = 0.1

2020-12-24 19:02:53,314 Training ended since there were no improvements in the last
learning rate step: 0.000240

2020-12-24 19:02:53,316 Best validation result at step 22000: 21.66 eval_metric.

2020-12-24 19:03:33,667 -----

2020-12-24 19:03:33,667 [DEV] partition [RECOGNITION] experiment [BW]: 1

2020-12-24 19:03:59,637 finished in 25.9699s

2020-12-24 19:03:59,637 *****
2020-12-24 19:03:59,637 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 1
 WER 27.36 (DEL: 11.34, INS: 4.54, SUB: 11.48)
2020-12-24 19:03:59,637 *****
2020-12-24 19:03:59,637 -----
2020-12-24 19:03:59,638 [DEV] partition [RECOGNITION] experiment [BW]: 2
2020-12-24 19:04:44,158 finished in 44.5203s
2020-12-24 19:04:44,159 *****
2020-12-24 19:04:44,159 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 2
 WER 26.85 (DEL: 10.62, INS: 4.70, SUB: 11.53)
2020-12-24 19:04:44,159 *****
2020-12-24 19:04:44,159 -----
2020-12-24 19:04:44,159 [DEV] partition [RECOGNITION] experiment [BW]: 3
2020-12-24 19:05:51,512 finished in 67.3533s
2020-12-24 19:05:51,513 *****
2020-12-24 19:05:51,513 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 3
 WER 26.77 (DEL: 10.44, INS: 4.62, SUB: 11.72)
2020-12-24 19:05:51,513 *****
2020-12-24 19:05:51,513 -----
2020-12-24 19:05:51,513 [DEV] partition [RECOGNITION] experiment [BW]: 4
2020-12-24 19:07:21,544 finished in 90.0310s
2020-12-24 19:07:21,545 *****
2020-12-24 19:07:21,545 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 4
 WER 26.74 (DEL: 10.41, INS: 4.62, SUB: 11.72)
2020-12-24 19:07:21,545 *****
2020-12-24 19:07:21,545 -----
2020-12-24 19:07:21,545 [DEV] partition [RECOGNITION] experiment [BW]: 5
2020-12-24 19:09:13,307 finished in 111.7622s
2020-12-24 19:09:13,308 *****
2020-12-24 19:09:13,308 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 5
 WER 26.69 (DEL: 10.38, INS: 4.62, SUB: 11.69)
2020-12-24 19:09:13,308 *****
2020-12-24 19:09:13,308 -----
2020-12-24 19:09:13,308 [DEV] partition [RECOGNITION] experiment [BW]: 6
2020-12-24 19:11:28,438 finished in 135.1299s
2020-12-24 19:11:28,438 -----
2020-12-24 19:11:28,439 [DEV] partition [RECOGNITION] experiment [BW]: 7
2020-12-24 19:14:06,398 finished in 157.9589s
2020-12-24 19:14:06,398 -----

2020-12-24 19:14:06,398 [DEV] partition [RECOGNITION] experiment [BW]: 8
2020-12-24 19:17:07,323 finished in 180.9245s
2020-12-24 19:17:07,323 -----
2020-12-24 19:17:07,323 [DEV] partition [RECOGNITION] experiment [BW]: 9
2020-12-24 19:20:30,357 finished in 203.0337s
2020-12-24 19:20:30,357 -----
2020-12-24 19:20:30,358 [DEV] partition [RECOGNITION] experiment [BW]: 10
2020-12-24 19:24:15,754 finished in 225.3967s
2020-12-24 19:24:15,755
=====

2020-12-24 19:24:40,785 [DEV] partition [Translation] results:
New Best Translation Beam Size: 1 and Alpha: -1
BLEU-4 21.66 (BLEU-1: 46.97, BLEU-2: 34.05, BLEU-3: 26.53,
BLEU-4: 21.66)
CHRF 44.19 ROUGE 47.04
2020-12-24 19:24:40,785 -----

2020-12-24 19:28:26,934 [DEV] partition [Translation] results:
New Best Translation Beam Size: 2 and Alpha: 1
BLEU-4 21.72 (BLEU-1: 46.65, BLEU-2: 33.86, BLEU-3: 26.44,
BLEU-4: 21.72)
CHRF 43.91 ROUGE 47.59
2020-12-24 19:28:26,934 -----

2020-12-24 19:28:52,758 [DEV] partition [Translation] results:
New Best Translation Beam Size: 2 and Alpha: 2
BLEU-4 21.76 (BLEU-1: 46.70, BLEU-2: 33.91, BLEU-3: 26.48,
BLEU-4: 21.76)
CHRF 43.94 ROUGE 47.59
2020-12-24 19:28:52,758 -----

2020-12-24 19:58:35,119 [DEV] partition [Translation] results:
New Best Translation Beam Size: 10 and Alpha: 2
BLEU-4 21.79 (BLEU-1: 45.54, BLEU-2: 33.47, BLEU-3: 26.35,
BLEU-4: 21.79)
CHRF 43.54 ROUGE 47.35
2020-12-24 19:58:35,120 -----

2020-12-24 20:02:27,801 [DEV] partition [Translation] results:
New Best Translation Beam Size: 11 and Alpha: 1
BLEU-4 21.81 (BLEU-1: 45.13, BLEU-2: 33.27, BLEU-3: 26.28,
BLEU-4: 21.81)
CHRF 43.60 ROUGE 47.62
2020-12-24 20:02:27,801 -----

2020-12-24 20:03:07,057 [DEV] partition [Translation] results:
New Best Translation Beam Size: 11 and Alpha: 2
BLEU-4 22.08 (BLEU-1: 45.73, BLEU-2: 33.74, BLEU-3: 26.64,
BLEU-4: 22.08)

CHRF 43.73 ROUGE 47.52
 2020-12-24 20:03:07,058 -----
 2020-12-24 20:07:10,017 [DEV] partition [Translation] results:
 New Best Translation Beam Size: 12 and Alpha: 1
 BLEU-4 22.14 (BLEU-1: 45.27, BLEU-2: 33.58, BLEU-3: 26.63,
 BLEU-4: 22.14)
 CHRF 43.84 ROUGE 47.86
 2020-12-24 20:07:10,017 -----
 2020-12-24 20:07:51,506 [DEV] partition [Translation] results:
 New Best Translation Beam Size: 12 and Alpha: 2
 BLEU-4 22.45 (BLEU-1: 46.04, BLEU-2: 34.17, BLEU-3: 27.07,
 BLEU-4: 22.45)
 CHRF 44.07 ROUGE 47.72
 2020-12-24 20:07:51,506 -----
 2020-12-24 20:09:55,546 *****
 2020-12-24 20:09:55,547 [DEV] partition [Recognition & Translation] results:
 Best CTC Decode Beam Size: 5
 Best Translation Beam Size: 12 and Alpha: 2
 WER 26.69 (DEL: 10.38, INS: 4.62, SUB: 11.69)
 BLEU-4 22.45 (BLEU-1: 46.04, BLEU-2: 34.17, BLEU-3: 27.07,
 BLEU-4: 22.45)
 CHRF 44.07 ROUGE 47.72
 2020-12-24 20:09:55,547 *****
 2020-12-24 20:12:21,556 [TEST] partition [Recognition & Translation] results:
 Best CTC Decode Beam Size: 5
 Best Translation Beam Size: 12 and Alpha: 2
 WER 27.78 (DEL: 9.09, INS: 6.06, SUB: 12.63)
 BLEU-4 22.06 (BLEU-1: 45.59, BLEU-2: 33.92, BLEU-3: 26.73,
 BLEU-4: 22.06)
 CHRF 44.14 ROUGE 46.99
 2020-12-24 20:12:21,556 *****

EXPS____4:

hyperparameters:
 lr_ctc_recog = 5.0
 lr_dec_split_recog = 0.7
 lr_dec_split_trans_de_en = 0.0
 num_layers_enc=num_layers_dec=3
 dropout = 0.1

2020-12-25 00:06:16,222 Training ended since there were no improvements in the last learning rate step: 0.000058
2020-12-25 00:06:16,224 Best validation result at step 44200: 22.03 eval_metric.
2020-12-25 00:07:03,876 -----
2020-12-25 00:07:03,876 [DEV] partition [RECOGNITION] experiment [BW]: 1
2020-12-25 00:07:29,232 finished in 25.3558s
2020-12-25 00:07:29,233 *****
2020-12-25 00:07:29,233 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 1
 WER 28.42 (DEL: 11.05, INS: 5.44, SUB: 11.93)
2020-12-25 00:07:29,233 *****
2020-12-25 00:07:29,233 -----
2020-12-25 00:07:29,233 [DEV] partition [RECOGNITION] experiment [BW]: 2
2020-12-25 00:08:18,865 finished in 49.6315s
2020-12-25 00:08:18,865 *****
2020-12-25 00:08:18,865 [DEV] partition [RECOGNITION] results:
 New Best CTC Decode Beam Size: 2
 WER 28.34 (DEL: 10.70, INS: 5.60, SUB: 12.04)
2020-12-25 00:08:18,865 *****
2020-12-25 00:08:18,865 -----
2020-12-25 00:08:18,865 [DEV] partition [RECOGNITION] experiment [BW]: 3
2020-12-25 00:09:34,576 finished in 75.7103s
2020-12-25 00:09:34,576 -----
2020-12-25 00:09:34,576 [DEV] partition [RECOGNITION] experiment [BW]: 4
2020-12-25 00:11:16,053 finished in 101.4761s
2020-12-25 00:11:16,053 -----
2020-12-25 00:11:16,053 [DEV] partition [RECOGNITION] experiment [BW]: 5
2020-12-25 00:13:24,026 finished in 127.9727s
2020-12-25 00:13:24,027 -----
2020-12-25 00:13:24,027 [DEV] partition [RECOGNITION] experiment [BW]: 6
2020-12-25 00:15:58,362 finished in 154.3353s
2020-12-25 00:15:58,363 -----
2020-12-25 00:15:58,363 [DEV] partition [RECOGNITION] experiment [BW]: 7
2020-12-25 00:18:59,195 finished in 180.8326s
2020-12-25 00:18:59,196 -----
2020-12-25 00:18:59,196 [DEV] partition [RECOGNITION] experiment [BW]: 8
2020-12-25 00:22:28,890 finished in 209.6938s
2020-12-25 00:22:28,891 -----
2020-12-25 00:22:28,891 [DEV] partition [RECOGNITION] experiment [BW]: 9
2020-12-25 00:26:25,030 finished in 236.1396s
2020-12-25 00:26:25,031 -----

2020-12-25 00:26:25,031 [DEV] partition [RECOGNITION] experiment [BW]: 10

2020-12-25 00:30:25,958 finished in 240.9269s

2020-12-25 00:30:25,958

=====

2020-12-25 00:30:45,996 [DEV] partition [Translation] results:

New Best Translation Beam Size: 1 and Alpha: -1

BLEU-4 22.03 (BLEU-1: 48.19, BLEU-2: 34.87, BLEU-3: 26.97,
BLEU-4: 22.03)

CHRF 45.32 ROUGE 47.39

2020-12-25 00:30:45,997 -----

2020-12-25 00:33:08,559 [DEV] partition [Translation] results:

New Best Translation Beam Size: 2 and Alpha: -1

BLEU-4 22.49 (BLEU-1: 48.37, BLEU-2: 35.26, BLEU-3: 27.46,
BLEU-4: 22.49)

CHRF 45.58 ROUGE 48.35

2020-12-25 00:33:08,559 -----

2020-12-25 00:33:50,970 [DEV] partition [Translation] results:

New Best Translation Beam Size: 2 and Alpha: 1

BLEU-4 22.67 (BLEU-1: 48.61, BLEU-2: 35.47, BLEU-3: 27.65,
BLEU-4: 22.67)

CHRF 45.76 ROUGE 48.45

2020-12-25 00:33:50,970 -----

2020-12-25 00:35:39,452 [DEV] partition [Translation] results:

New Best Translation Beam Size: 3 and Alpha: -1

BLEU-4 22.76 (BLEU-1: 47.57, BLEU-2: 35.04, BLEU-3: 27.57,
BLEU-4: 22.76)

CHRF 45.34 ROUGE 48.53

2020-12-25 00:35:39,452 -----

2020-12-25 00:36:24,584 [DEV] partition [Translation] results:

New Best Translation Beam Size: 3 and Alpha: 1

BLEU-4 22.95 (BLEU-1: 48.04, BLEU-2: 35.38, BLEU-3: 27.83,
BLEU-4: 22.95)

CHRF 45.58 ROUGE 48.65

2020-12-25 00:36:24,585 -----

2020-12-25 00:42:00,695 [DEV] partition [Translation] results:

New Best Translation Beam Size: 5 and Alpha: 1

BLEU-4 23.05 (BLEU-1: 48.01, BLEU-2: 35.46, BLEU-3: 27.94,
BLEU-4: 23.05)

CHRF 45.76 ROUGE 48.89

2020-12-25 00:42:00,695 -----

2020-12-25 01:11:01,603 *****

2020-12-25 01:11:01,603 [DEV] partition [Recognition & Translation] results:

Best CTC Decode Beam Size: 2

Best Translation Beam Size: 5 and Alpha: 1

WER 28.34 (DEL: 10.70, INS: 5.60, SUB: 12.04)
BLEU-4 23.05 (BLEU-1: 48.01, BLEU-2: 35.46, BLEU-3: 27.94,
BLEU-4: 23.05)
CHRF 45.76 ROUGE 48.89
2020-12-25 01:11:01,603 *****
2020-12-25 01:12:04,092 [TEST] partition [Recognition & Translation] results:
Best CTC Decode Beam Size: 2
Best Translation Beam Size: 5 and Alpha: 1
WER 27.92 (DEL: 8.66, INS: 5.61, SUB: 13.64)
BLEU-4 22.59 (BLEU-1: 47.59, BLEU-2: 35.16, BLEU-3: 27.60,
BLEU-4: 22.59)
CHRF 45.58 ROUGE 48.82
2020-12-25 01:12:04,092 *****