



## 2020 하계 계절학기 집중이수제 주간학습보고서 (1주차)

이름	최 용호	학과(전공)	지능기전공학부																														
학번	17013253	학년	4																														
연락처		e-Mail																															
과목명	자기주도 창의전공 I	신청학점	3	분반	12																												
학습기간	2020.06.29 ~ 2020.07.05	학습주차	1	학습시간	50																												
창의과제	저화질 영상에 대한 Video Captioning 네트워크 성능 향상 연구																																
금주 학습목표	Video Captioning 논문 조사 및 CNN 학습 및 점수 비교 ( 다양한 최신 CNN 네트워크 적용 및 성능 평가 비교 )																																
학습내용	S2VT의 Train / Test 의 Sequence 및 입출력의 shape 파악. MSVD CNN 피쳐들을 뽑아내는 방법 숙지. S2VT에 COCO 평가점수 코드 결합																																
참고자료 및 문헌	S2VT : Sequence to Sequence Video to Text ( <a href="https://vsubhashini.github.io/s2vt.html">https://vsubhashini.github.io/s2vt.html</a> ) Microsoft COCO Captions: Data Collection and Evaluation Server ( <a href="https://arxiv.org/pdf/1504.00325.pdf">https://arxiv.org/pdf/1504.00325.pdf</a> ) Reconstruction Network for Video Captioning ( <a href="https://arxiv.org/pdf/1803.11438.pdf">https://arxiv.org/pdf/1803.11438.pdf</a> )																																
학습방법	위 저화질에 대한 Video Captioning 네트워크 성능 향상 연구에 앞서서, 고화질 영상에서의 Video Captioning에 대한 연구의 점수를 확보해야 저화질에서 Video Captioning 성능이 어느 정도 낮아지는지 알 수 있다. 따라서, 각 논문에서 실험했던 VGG16, InceptionV4 모델에서 MSVD 데이터에서 피쳐를 직접 뽑아 S2VT모델을 통해 논문 점수를 원복시켜보았다.																																
학습성과 목표달성도	<div>실험 환경</div> <div>CPU : Intel(R) Core(TM) i9-9900X GPU <u>Geforce RTX 2080 Ti (12GB) 3EA</u></div> <div>Python : 2.7 Caffe, CUDA : 8.0</div> <div><table><tr><th>Model</th><th>BLEU-4</th><th>METEOR</th><th>ROUGE-L</th><th>CIDEr</th></tr><tr><td>S2VT (Inception-V4)</td><td>39.6</td><td>31.2</td><td>67.5</td><td>66.7</td></tr></table></div> <div><table><tr><td>S2VT (ours)</td><td></td><td></td></tr><tr><td>- Flow (AlexNet)</td><td>24.3</td><td>(8)</td></tr><tr><td>- RGB (AlexNet)</td><td>27.9</td><td>(9)</td></tr><tr><td>- RGB (VGG) random frame order</td><td>28.2</td><td>(10)</td></tr><tr><td>- RGB (VGG)</td><td>29.2</td><td>(11)</td></tr><tr><td>- RGB (VGG) + Flow (AlexNet)</td><td>29.8</td><td>(12)</td></tr></table></div> <div><div>Extracted VGG</div><div>CIDEr: 0.506 Bleu_4: 0.371 Bleu_3: 0.485 Bleu_2: 0.599 Bleu_1: 0.741 ROUGE L: 0.652 METEOR: 0.295</div></div> <div><div>Extracted_InceptionV4</div><div><ul style="list-style-type: none"><li>CIDEr: 69.1</li><li>bleu_4: 41.7</li><li>bleu_3: 52.5</li><li>bleu_2: 62.9</li><li>bleu_1: 76.5</li><li>Rouge L: 67.4</li><li>METEOR: 31.5</li></ul></div></div>					Model	BLEU-4	METEOR	ROUGE-L	CIDEr	S2VT (Inception-V4)	39.6	31.2	67.5	66.7	S2VT (ours)			- Flow (AlexNet)	24.3	(8)	- RGB (AlexNet)	27.9	(9)	- RGB (VGG) random frame order	28.2	(10)	- RGB (VGG)	29.2	(11)	- RGB (VGG) + Flow (AlexNet)	29.8	(12)
Model	BLEU-4	METEOR	ROUGE-L	CIDEr																													
S2VT (Inception-V4)	39.6	31.2	67.5	66.7																													
S2VT (ours)																																	
- Flow (AlexNet)	24.3	(8)																															
- RGB (AlexNet)	27.9	(9)																															
- RGB (VGG) random frame order	28.2	(10)																															
- RGB (VGG)	29.2	(11)																															
- RGB (VGG) + Flow (AlexNet)	29.8	(12)																															
내주 계획	저화질 영상 생성 알고리즘 개발과 저화질 영상에 기존의 Video Captioning 네트워크 적용 ( QP 조정을 통한 저화질 영상에 대한 적용 및 평가 )																																

2020. 07. 06.

지도교수

(인)