병렬프로그래밍 7주차

팀명: TEAM\_C\_TENCORE

팀원: 소프트웨어융합대학 20143148 이호중

소프트웨어융합대학 20135357 한승탁

컴퓨터 공학과 20135114 김용화

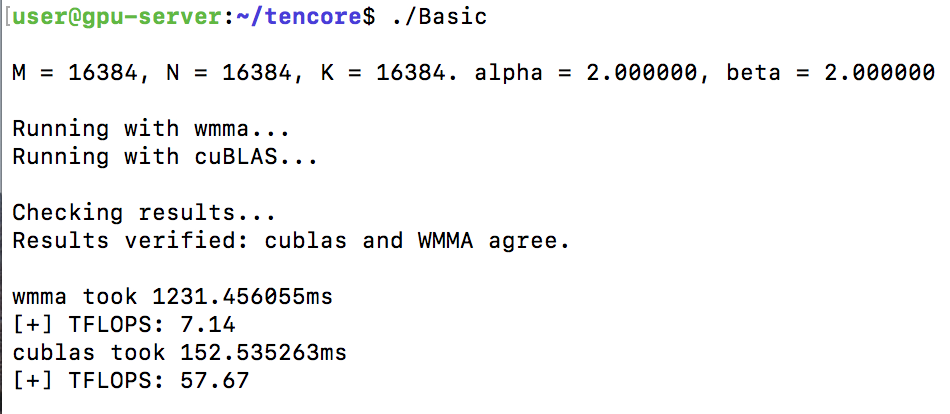
콘텐츠IT 20125345 조준형

목표: TENSOR CORE 기반의 행렬곱 가속화 및 딥러닝 적용 방식 분석

기간:2018.09.01 ~ 2018.12.20

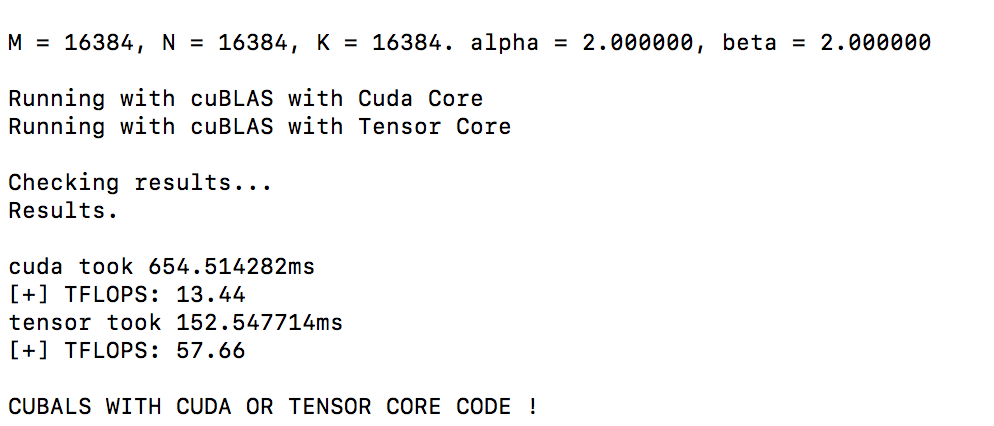
대학생 팀 발표 내용:

텐서코어를 이용하여 속도 측정하였습니다.

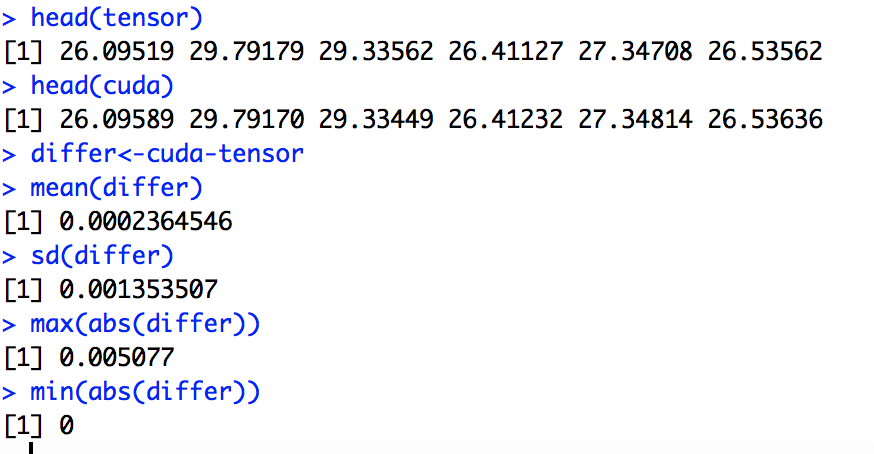


Wmma를 이용한 행렬연산과 cublas를 이용하여 tflops를 측정하였습니다.

두 코드 모두 tensor코어에서 동작하였고 wmma비교하여 cublas는 최적화가 되어있기 때문에 큰 속도 차이를 보였습니다.

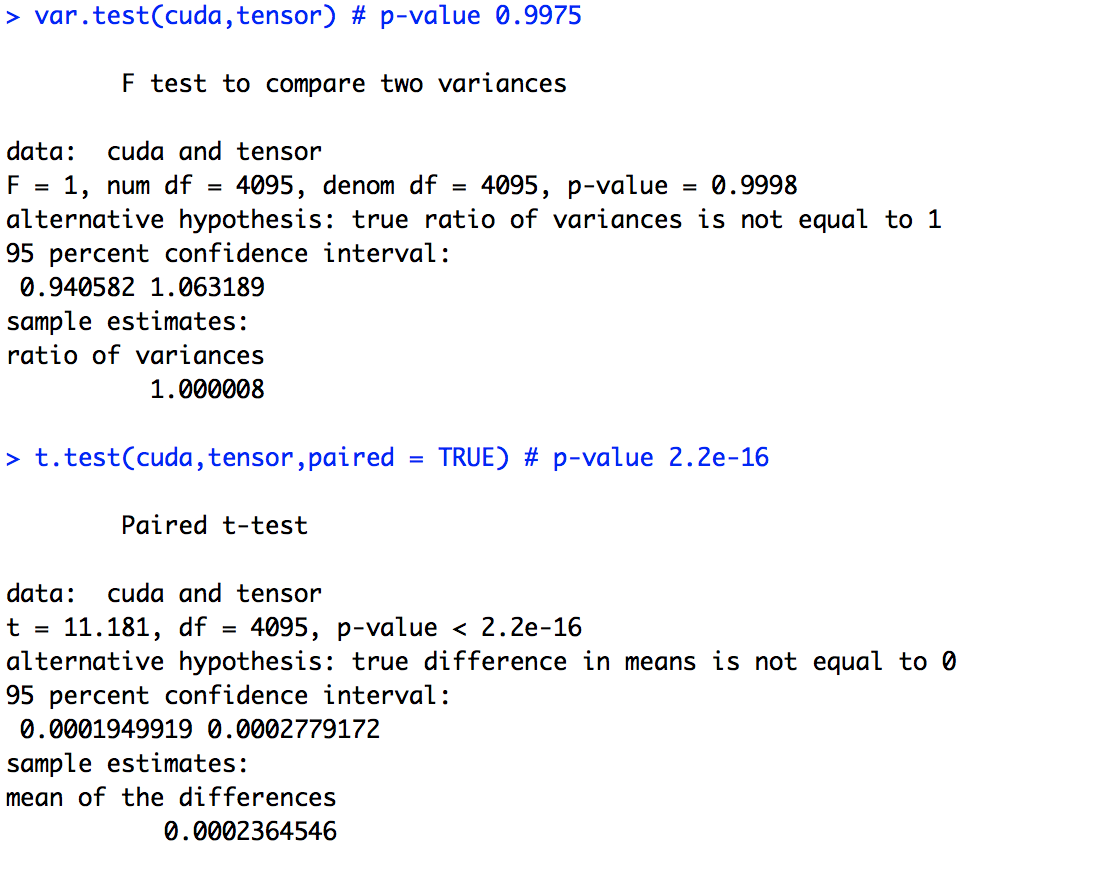


Cublas에 옵션을 변경하여 같은 행렬을 하나는 쿠다 코어를 이용하여 나머지 하나는 tensor 코어를 이용하여 연산하였을 때에 결과입니다. cuda코어를 이용하였을 때에는 13.44 tlops이며 텐서코어를 이용하였을 때에는 57.66 tflops 로 약4배의 연산 성능 차이를 보였습니다.



추가적으로 쿠다코어는 32비트 연산을 하고 텐서코어는 16비트 연산을 하고 32비트로 누적하기 때문에 코어별 연산결과에 차이가 있는지 알아보기 위하여 결과값을 데이터셋으로 만들어 둘의 결과에 차이가 있는 지를 비교하였습니다. 그 결과 평균적으로 0.002364546정도의 차이를 보이며

표준편차는 0.001353507 이었습니다. 가장 많이 차이가 나는 결과 값은 0.005077이었으며 차이가 나지 않는 경우도 있었습니다.



이 값의 차이가 통계적으로 유의한지 테스트를 하였고 그 결과 p-value가 0.05이하로

쿠다코와 텐서코어의 연산 결과 값은 통계적으로 차이가 있다는 것을 확인하였습니다.

미팅 내용:

대학원생팀에서 Nvprof를 이용하여 코어 동작 시간을 확인 할 수 있으니 참고해보라고 조언해주셧습니다. 그리고 cuda코어와 텐서코어 비교하여 최종 보고서 작성을 시작할 것이라고 말씀드렸고 cpu와 strassen방식은 결과값이 나오는 대로 추가하기로 하였습니다.

미팅 사진:

