I am a reviewer for I will Begin by giving a summary this paper presents an A
and an algorithm on TPU and analyzes the memory and instruction bandwidth of a n
n algorithms I will now go over some struts I will now go over some strengths
first of all the problem is well motivated second the related work is
comprehensively studied and discussed and the entire story is easy to follow for
readers and lastly the methodology to study the algorithm from the hardware
bottlenecks is interesting here are some weaknesses I suggest the authors to
include a brief discussion of TPU for example what it can be done and what I
cannot be done efficiently to experimental evaluation is a bit problematic as
well how about other methods on gpus gpus like hashing and graph-based methods
beside the f a i s s Baseline and a figure 3 the highest recall shown for glove
is 0.9 is there a recall limitation for the proposed algorithm and finally the
technical contribution of the by stage partial reduction and scoring is limited