note of《FFT Convolutions are Faster than Winograd on Modern CPUs, Here's Why》

zxp

October 12, 2024

1 content

《FFT Convolutions are Faster than Winograd on Modern CPUs, Here's Why》 is an 18-year paper detailing the pros and cons of Winograd and FFT.

Winograd is generally considered to be faster than FFT in 18 years because Winograd is computed less frequently. The paper points out that the reason why Winorad and FFT are used to accelerate convolution is that the calculation of the result requires less computation. but the paper believes that the AI of FFT is greater, one is because FFT is calculated in the complex field, and the other is because the data used in FFT calculation is more dense than Winograd. Thus FFT computes convolution no slower than Winograd. roofline was used in the analysis of the paper, and the content related to rooflin can be referred to this paper.

The Wingograd library used in the paper experiment is the work of others, the work of the paper is to optimize FFT, optimize two kinds, basic and Gaussian FFT, Gaussian FFT uses Gaussian way to calculate the complex field, the number of times a complex multiplication operation is calculated from 4 times to 3 times.

The way of selecting the test benchmark can also be refered.

2 Feelings

This paper is of great value. The introduction part of this paper has a detailed review of the history of Winograd. Please refer to the part related to Winograd here.