

Positional Convolution Experts

Idea

Un'architettura che introduce esperti convoluzionali specializzati per posizione.

La PCE divide l'immagine in patch e utilizza un meccanismo di routing posizionale per indirizzare ciascuna patch verso esperti specializzati.

Architettura PCE

- Divisione in patch dell'immagine di input
- Sistema di routing che combina informazioni di contenuto e posizione
- Esperti specializzati per diverse regioni/caratteristiche
- Combinazione finale delle feature map per risultati arricchiti

Metodologia di Training (3 Fasi)

- Fase 1: Stabilizzazione
Tre approcci alternativi per inizializzare il sistema:
A) Routing deterministico: con rumore controllato
B) Cosine similarity: con chiavi EMA inizializzate tramite K-Means
C) Distribuzione uniforme: tra gli esperti
- Fase 2:
Viene scelto un meccanismo di routing: MLP, KeyAttention con MAE, Gubme Softmax

Valutazione e Metriche

Dataset: CIFAR-10, Tiny-ImageNet, Pascal VOC, Camelyon

Task:

- Classificazione: Top-1/Top-5 accuracy
- Object Detection**: mAP (Mean Average Precision)
- Segmentazione: mIoU (Mean Intersection over Union)

Risultati attesi

Ci si aspetta un miglioramento delle performance rispetto a CNN classiche grazie agli esperti e al routing posizionale. La rete risultarebbe più interpretabile, grazie alla possibilità di analizzare le scelte del router e la distribuzione degli esperti.

Un possibile pareggio con architetture di CNN più moderne. Curioso il confronto con DyNet da cui ho preso spunto.

Il costo complessivo della rete dovrebbe essere ridotto rispetto alle CNN classiche, questo grazie agli esperti, non tutti i parametri saranno attivi, dovrebbe esserci un costo minore, in oltre l'utilizzo di MAE che non richiede gradienti potrebbe migliorare ulteriormente il training