

DynamicViT f/N yÍšØeHv,,%oÆ%oÉ Transformer j!W<ÿ • •ÇRˆˆ NârLz u•S ÿ Dynamic Token Sparsification

1. h8_ÃR e°

- **Rˆˆ NârLRjg•**ÿ • •Ç•{‘İ~\$~,,mKj!WWÿ Prediction Moduleÿ • \B<ÃO0kĬN*tokenv,,‘Í%o•`'R epÿ • keN" _ C
- **xINöSËY}v,,z u•`'***ÿ R)u(•êlèa R>g:R6v,,ry`'ÿ SsO•Rjg•T v,,tokenR ^ N %oÃR ÿ NÍÉýšØeH^v^L{¡{—0
- **lèa R>c©x {Vue**ÿ <~Ãeö• •ÇGumbel-SoftmaxTØElèa R>c©x ÿ Attention Maskingÿ [žs°zİR0zİO S ÿ

2. N;%o•O R¿

- **šØeH`'***ÿ W(ImageNetN ÿ FLOPsQİ\ 31%~37%ÿ T T ‘İcÐSG40%+ÿ |¾^!c_Y1NÅ0.5%NåQ...ÿ Y,D
- **• u(`'***ÿ • ‘MY yÍViTSØOSÿ Y,DeiT0 LV-ViTÿ ÿ e/c N T •“QeR •ˆs‡ÿ Y,384×384ÿ 0
- **Sİ%oã‘É`'***ÿ Sİ%oÆS f>y:j!W• ke∈Zq&NŽVpPİQs•.S:Wßÿ Y,Vp5ÿ 0

3. b∈g/~Æ,,

- **R \BRjg•**ÿ R 3N*–6κμRjg•ÿ Y,OÝc kÔO<c Ó ãr #Á³"H34%RiOYtokenÿ 0
- **c_Y1Qýep**ÿ ~ÔT R |{c_Y10 „,™•c_Y1ÿ j!NÿSÿYËj!W<^LN:ÿ TØEkÔO<~!g_c_Y1ÿ c§R6Rjg•s‡ÿ 0
- **c`'t R • **ÿ vîc¥c R epOÝuYTop-k tokenÿ eà— ∈ †QSi_@R `0

4. [žšOE~Ógœ

- **j!W<[ùkÔ**ÿ DynamicViT-LV-S/0.7NÅ— 4.6 GFLOPsÿ |¾^!83.0%ÿ O NŽEfficientNet{ICNNj!W<ÿ ^h2ÿ
- **bìU`'***ÿ W(Y'j!W<ÿ Y,DeiT-Bÿ TØEY'•“Qeÿ 384×384ÿ N NÍg eHÿ ^h5ÿ 0

5. _n•N a NI

- Nãx]ò_ n•ÿ <https://github.com/raoyongming/DynamicViT>
- **•!s. **ÿ N:ViTR • _ •ÝN†e°e¹T ÿ Rˆˆ zz•ôz u•`'ÿ ÿ g*geSİbìU•ó%oÆ~'R |{{INûR¡0

Y,— fô<æ~ÆEv,,[žs°~Æ,,ÿ Y,~,mKj!WW~Óg,,b <~Ã...SÂÿ ÿ Sİ•ÛN kegâ– <°e‡b Nãx ^“0