

•Û{Ç°e‡0 DynamicViT: Efficient Vision Transformers with Dynamic Token Sparsification0 cĐQúN†N yÍ
Transformerŷ ViTŷ j!W<0 NâN f/<°e‡v,,h8_ÃQ...[!i,•đŷ

1. **h8_Ã•ĩ~N R¨g.**

- **€İfo**ŷ %œÆ%ÉTransformer\ VpPİR RrN:Y N*NärLŷ tokenŷ •Ū^LY t ŷ OF<ı{—b g,šÖŷ VâN:b@g Nâ
- **%œÂ[ß**ŷ j!W<g ~Ê~,mKNÂO•VNŽ\ 'İQs•.NärLŷ QvOYNärLQ—OY0
- **vİh **ŷ R¨ ¸ÆR+^vRjg•Q—OYNärLŷ Qİ\ <ı{—'İŷ FLOPsŷ ^vcĐSGc¨t • ^İŷ T eöOÝc j!W<|¾^!0

2. **e!İÖR e°**

- **R¨ NärLz u•S **ŷ
- **R \BRjg•**ŷ W(Transformerv,,N T \B• keRjg•NärLŷ Y,12\Bj!W<N-W({,40 70 10\BRjg•ŷ 0
- **~,mKj!WW**ŷ •{İ~šj!WW<Ä0kİN*NärLv,,!%œ`'R epŷ ~ÓT \@•èry_•TœQh\@N N e‡u b Rjg•Q³{V0
- **!èa R'cœx **ŷ ••ÇGumbel-Softmax[žs°Si_®Rjg•ŷ ••QMvôc¥N"_ NärL[ü•ôv,,^v^L<ı{—•ĩ~ŷ T eö\O...

3. **b€g/~Æ,,**

- **<~_Ã{Vue**ŷ
- O•u(eY^ j!W<ŷ SŸYËViTŷ •Ū^Lwâ<Æ,,™•ŷ ~İg_Rjg•T j!W<v,,•“Qún eY^ N •ô0
- _ QekÔs‡c_Y1ŷ ratio lossŷ c§R6kİ\BRjg•kÔO<ŷ xnOYR • eHgœ0
- **c¨t O S **ŷ vôc¥c ~,,¾kÔO<OÝuYšØR NärLŷ xİNöSËY}v,,z u•~Óg,,[žs°[ž—ER • 0

4. **[žšœ~Ógœ**

- **'€ŷ**ŷ W(ImageNetN ŷ Rjg•66%NärLeöŷ
- **<ı{—'İ**ŷ Qİ\ 31%~37% FLOPs0
- **• ^İ**ŷ T T 'İcĐSG40%NâN 0
- **|¾^!c_Y1**ŷ NÂN —M0.5%NâQ...ŷ Y,DeiT-SNÎ79.8%—M•ó79.3%ŷ 0
- **[ùkÔSOTA**ŷ R¨ Rjg•O NŽ—Y` Rjg•Tœ~Óg,,S N 'Çh7ŷ •¾R0N šØeHCNNŷ Y,EfficientNetŷ vø_Sv,

5. **Sİ%œS N %œã'Ê`!**

- Rjg•Çz • n €Zq&NŽVpPİN-v,,Qs•.S:Wßŷ Y,riOSN-_Ãŷ ŷ Xž_:j!W<Sİ%œã'Ê`!0

6. **^”u(o\R>**

- SİbİU•ófôY'j!W<ŷ Y,DeiT-Bŷ TœšØR •s‡•“Qeŷ 384×384ŷ ŷ •ŪN kešœ<Áe!İÖv,,• u(`'0

`;~Ó

DynamicViT• •ÇR¨ NärLRjg•f>,,WcĐSGN†ViTv,,eHs‡ŷ N:[žeö%œÆ%ÉNûRıcĐO>N†Sİ^Lv,,O S •İ_,ŷ T e