

•Ü{Çˆºe‡0 DynamicViT: Efficient Vision Transformers with Dynamic Token Sparsification0 cĐQúN†N yÍ
Transformerj!Wˆ0 NâN f/ˆºe‡v„h8_ÃQ...[1ÿ

1. **h8_Ã•~**

- **€İfo**ÿ %œ%ÉTransformerÿ Y,ViT0 DeiTÿ ••Ç\ VpPİR RrN:Y N*NärLÿ tokenÿ •Ű^LY t ÿ OFç{—b ç
- **%Â[ß**ÿ ^v—^b@g NärL[ùg ~É~,mK•ýg T {l!s.ÿ ç,Y Q—OYNärLSiNâ^«yû—d€ N_qTÍQÆxn`0

2. **N;%•!s.**

- **Rˆˆ NärLz u•S **ÿ cĐQúN yÍ•{İ~\$~,mKj!WWÿ h9cn•“QeRˆˆ ‹ÄO0kİN*NärLv,„Í%•`'R epÿ • keRjg•Q—
- **\B~\$Rjg•**ÿ W(Transformerv„N T \BcÒQe~,mKj!WWÿ • keQİ\ NärLep`İÿ OˆY,ÿ OÝuYkÔOˆc QàOU
- **xİNöSÉY}**ÿ u1NŽ•êlèa Rˆv„ry`İÿ z u•S T v„NärLNÍe/c ^v^Lç{—ÿ [Ž—ER • eHgœf>„W0

3. **Qs•.b€g/**

- **~„mKj!WW**ÿ ~ÓT \@•èry_•ÿ SUN*NärLÿ TœQh\@ry_•ÿ VpPİN N e‡ÿ ~„mKOÝuY/N" _ i,s‡0
- **lèa Rˆc@x ÿ Attention Maskingÿ **ÿ ••Ç\O...=^«Rjg•NärLN QvNÖNärLv„NˆN'ÿ ••QMvôc¥R —d[ü•ôv
- **ç~—Ãvİh **ÿ ~ÓT R [ç_Y10 „,™•c_Y1ÿ N SÝYËj!Wç[üÿPÿ TœNärLkÔOˆ~!g_c_Y1ÿ zİR0ziO S 0

4. **[žšœ~Ógœ**

- **eHs‡cĐSG**ÿ W(ImageNetN ÿ Rjg•66%v„NärLÿ OÝuYkÔOˆ< Ó ã•đ—_oñ
- **ç{—İQİ\ **ÿ FLOPs—MON31%~37%0
- **c`t • ^İ**ÿ T T İcĐSG40%NâN 0
- **|¾^İc_Y1**ÿ R [QÆxns‡N —MNÂ0.2%~0.5%0
- **j!WçQ[1`**ÿ • u(NŽY yÍœ%ÉTransformerÿ Y,DeiT0 LV-ViTÿ ÿ N W(fôY'j!Wçÿ Y,384x384•“Qeÿ N M

5. **O RçN a NI**

- **Rˆˆ • ^”`**ÿ h9cn•“QeQ...[1Q³[šRjg•{Vueÿ O NŽ—Y` Rjg•b WGS N `Çh70
- **Sİ%œã‘Ê`**ÿ Sİ%œÆS f>y:j!Wç• n €Zq&NŽVpPİN-v„Qs•.S:Wßÿ Y,riOSN;OSÿ 0
- **bİU`**ÿ N:QvNÖ%œ%ÉNûRjÿ Y,hÂmK0 R Rrÿ v„R • cĐOˆN†eˆ` •İ0

6. **_ n•N •Dn•**

- Năx]ò_ n•ÿ [GitHub”pc¥](https://github.com/raoyongming/DynamicViT)0

{€€ Š NKÿ DynamicViT• •ÇRˆˆ Rjg•Q—OYNärLÿ W(QàNNN c_Y1|¾^İv„`ÂQµN f>„WcĐSGN†%œ%É