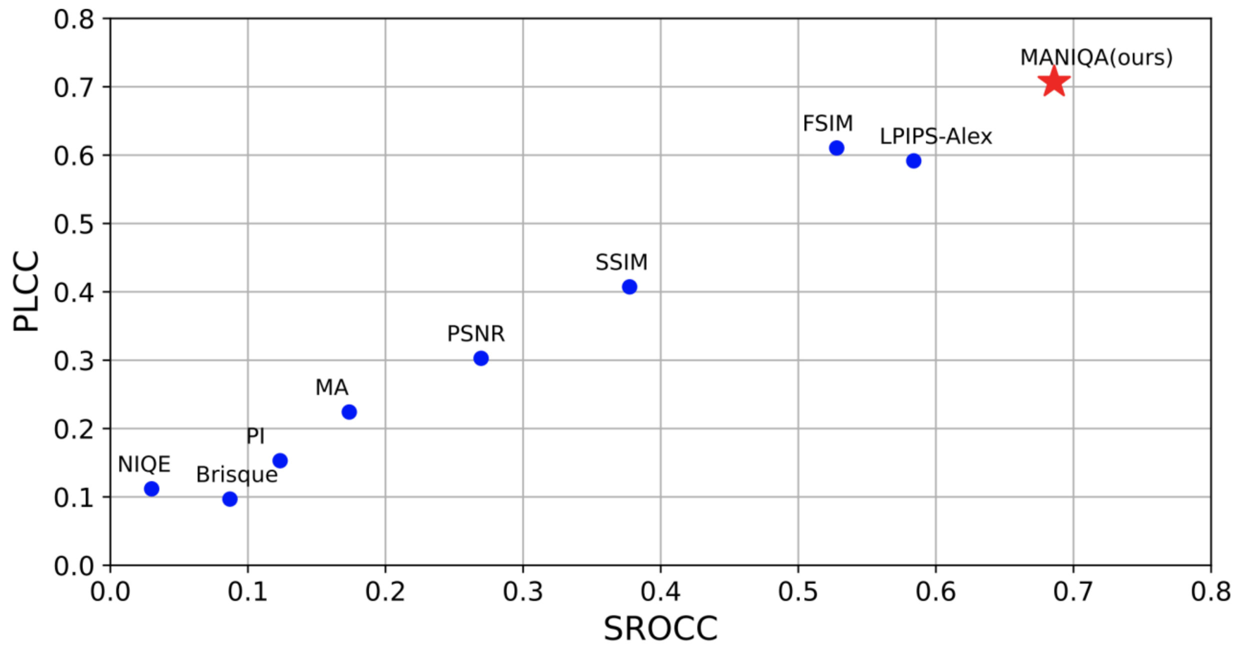
MANIQA: Multi-dimension Attention Network for No-Reference Image Quality Assessment

Sidi Yang* Tianhe Wu* Shuwei Shi Shanshan Lao Yuan Gong Mingdeng Cao Jiahao Wang Yujiu Yang†
Tsinghua Shenzhen International Graduate School, Tsinghua University

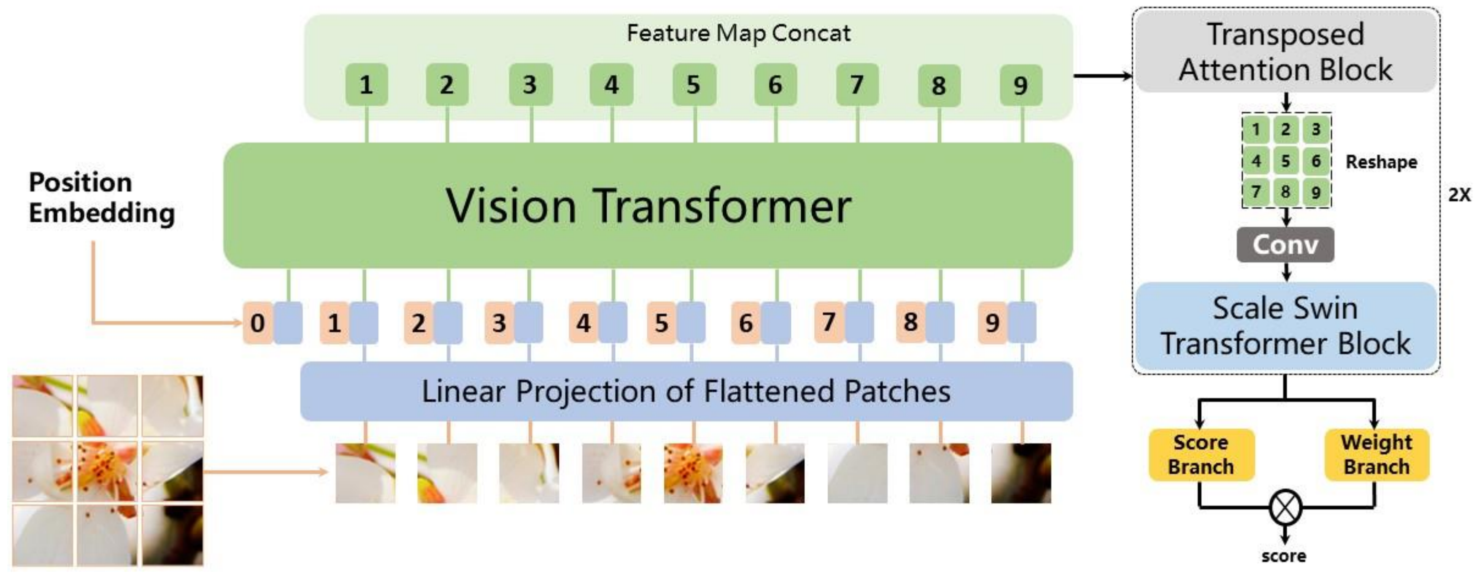
Section1: Introduction



Our contributions:

- We propose the Transposed Attention Block (TAB) and the Scale Swin Transformer Block (SSTB).
- We propose a dual branch structure for patch-weighted quality prediction.
- Our method ranked **first place** in the final testing phase of the NTIRE 2022 Perceptual Image Quality Assessment Challenge Track 2: No-Reference.

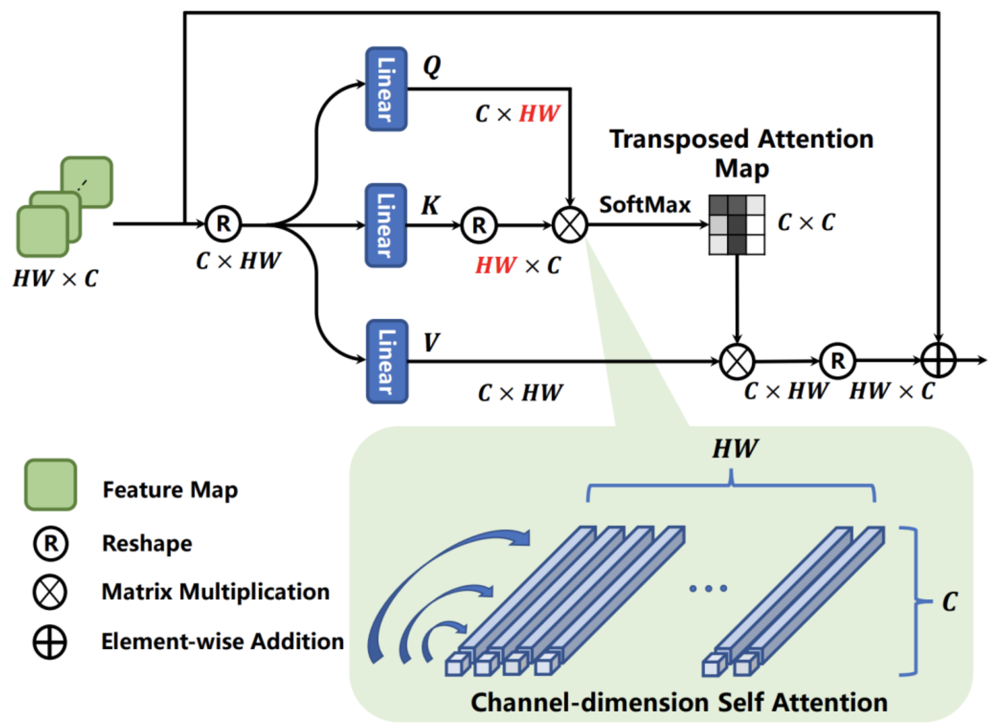
Section2: Model Architecture



The architecture of the proposed approach

Section3: Method

3.1 Transposed Attention Block

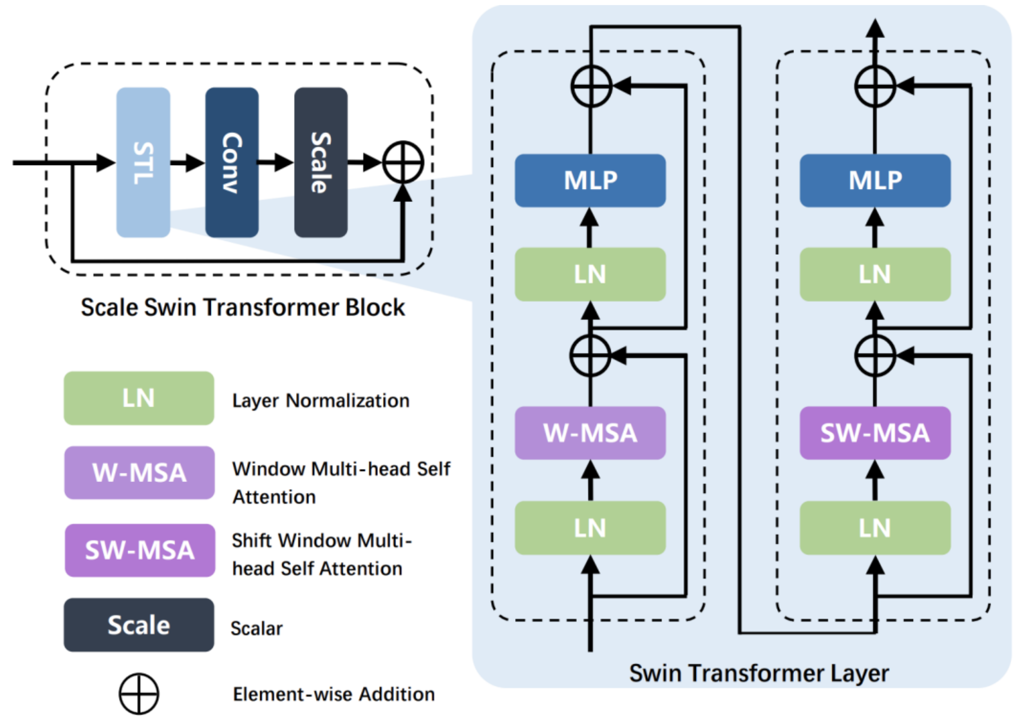


From the concatenated feature, our TAB plays:

$$\hat{X} = Attn(\hat{Q}, \hat{K}, \hat{V}) + X$$

$$Attn(\hat{Q}, \hat{K}, \hat{V}) = \hat{V} \cdot Softmax(\hat{Q} \cdot \hat{K} / \alpha)$$

3.2 Scale Swin Transformer Block



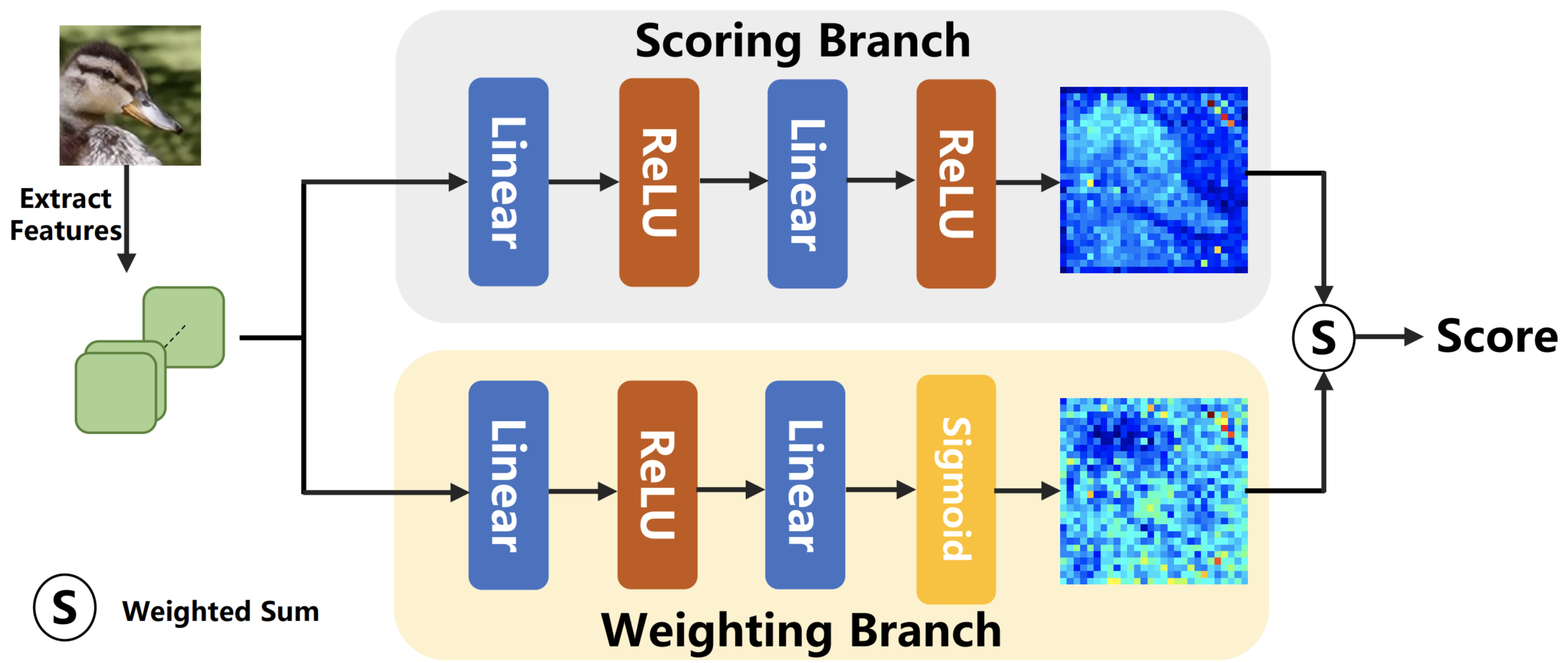
Given the input feature $F_{i,0}$:

$$F_{i,j} = H_{STL_{i,j}}(F_{i,j}), j = 1,2$$

The output of SSTB is formulated as:

$$F_{out} = \alpha \cdot H_{CONV}\left(H_{STL}(\hat{F}_{i,2})\right) + \hat{F}_{i,0}$$

3.3 Dual branch structure for patch-weighted quality prediction



Section4: Results

Comparison of MANIQA v.s. state-of-the-art NR-IQA algorithms

	LIVE		CSIQ		TID2013		KADID-10K	
	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
DIIVINE [34]	0.908	0.892	0.776	0.804	0.567	0.643	0.435	0.413
BRISQUE [28]	0.944	0.929	0.748	0.812	0.571	0.626	0.567	0.528
ILNIQE [55]	0.906	0.902	0.865	0.822	0.648	0.521	0.558	0.528
BIECON [16]	0.961	0.958	0.823	0.815	0.762	0.717	0.648	0.623
MEON [26]	0.955	0.951	0.864	0.852	0.824	0.808	0.691	0.604
WaDIQaM [3]	0.955	0.960	0.844	0.852	0.855	0.835	0.752	0.739
DBCNN [59]	0.971	0.968	0.959	0.946	0.865	0.816	0.856	0.851
TIQA [52]	0.965	0.949	0.838	0.825	0.858	0.846	0.855	0.850
MetaIQA [61]	0.959	0.960	0.908	0.899	0.868	0.856	0.775	0.762
P2P-BM [51]	0.958	0.959	0.902	0.899	0.856	0.862	0.849	0.840
HyperIQA [37]	0.966	0.962	0.942	0.923	0.858	0.840	0.845	0.852
TReS [9]	0.968	0.969	0.942	0.922	0.883	0.863	0.858	0.915
MANIQA(Ours)	**0.983**	**0.982**	**0.968**	**0.961**	**0.943**	**0.937**	**0.946**	**0.944**

Visualization