PanFlowNet: A Flow-Based Deep Network for Pan-sharpening

Gang Yang[1], Xiangyong Cao[2*], Wenzhe Xiao[2], Man Zhou[3], Aiping Liu[1*], Xun Chen[1], Deyu Meng[2]

1 University of Science and Technology of China; 2 Xi'an Jiaotong University; 3 Nanyang Technological University

ICCV23 PARIS

Task

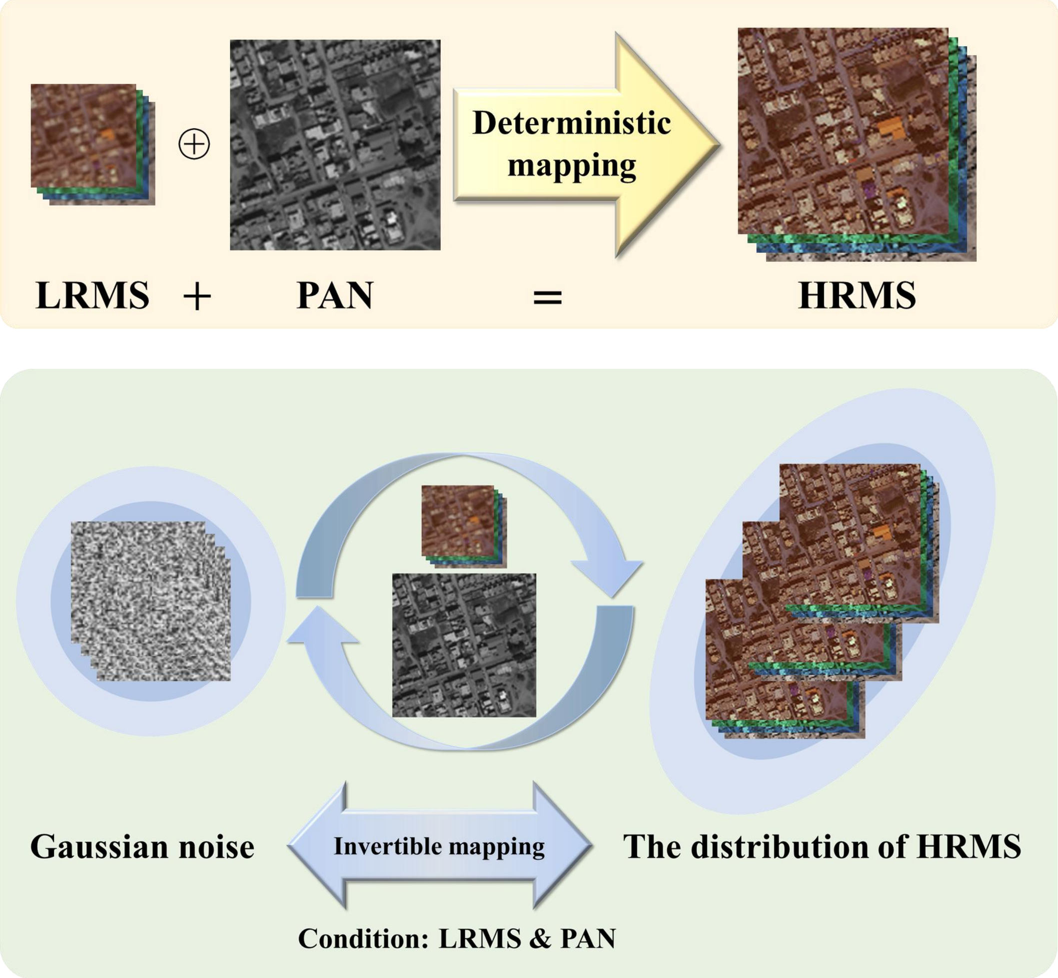
Pan-sharpening aims to generate a high-resolution multispectral (HRMS) image by integrating the spectral information of a low-resolution multispectral (LRMS) image with the texture details of a high-resolution panchromatic (PAN) image.

Motivation

- **ill-posed nature of the super-resolution (SR) task.**

we propose a flow-based pan-sharpening network (**PanFlowNet**) to directly learn the **conditional distribution** of HRMS image given LRMS image and PAN image instead of learning a deterministic mapping.

The proposed PanFlowNet can **generate diverse HRMSs** given the LRMS and PAN images as well as the Gaussian noise sample and thus can alleviate the ill-posed issue to some extent.



Probabilistic flow model

To alleviate the ill-posed issue, this work aims to learn the conditional distribution of the HRMS image $\boldsymbol{H}$, given the LRMS image $\mathbf{L}$ and the PAN image $\mathbf{P}$. $(\boldsymbol{P}_{\boldsymbol{H}|\mathbf{L},\boldsymbol{P}}(\boldsymbol{H} \mid \boldsymbol{L}, \boldsymbol{P};\ \boldsymbol{\theta}))$

In this conditional setting, $f_{\boldsymbol{\theta}}(\cdot)$ can map a LRMS-PAN-HRMS image pair to a latent variable z

$$\mathbf{z} = f_{\boldsymbol{\theta}}(\mathbf{H};\mathbf{L},\mathbf{P}), \longrightarrow \mathbf{H} = f_{\boldsymbol{\theta}}^{-1}(\mathbf{z};\mathbf{L},\mathbf{P}),\ \ \mathbf{z} \sim P_{\mathbf{z}}(\cdot),$$

the probability density function $P_{H|L,P}(H \mid L, P;\ \theta)$ can be accurately defined by using the change-of-variables formula

$$P_{\mathbf{H}|\mathbf{L},\mathbf{P}}(\mathbf{H}|\mathbf{L},\mathbf{P};\boldsymbol{\theta}) = P_{\mathbf{z}}(f_{\boldsymbol{\theta}}(\mathbf{H};\mathbf{L},\mathbf{P}))\left|\det\frac{\partial f_{\boldsymbol{\theta}}(\mathbf{H};\mathbf{L},\mathbf{P})}{\partial \mathbf{H}}\right|,$$

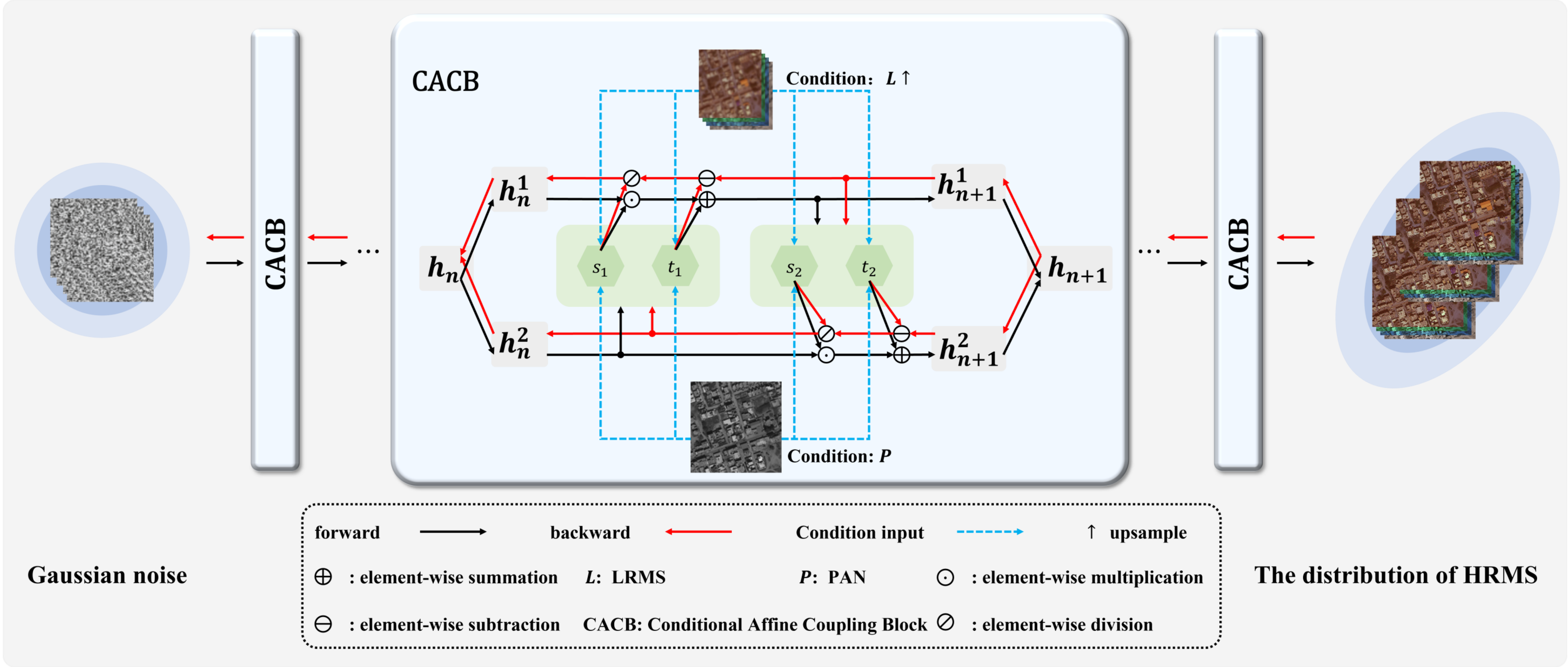
$$\mathcal{L}(\boldsymbol{\theta};\mathbf{L},\mathbf{P},\mathbf{H}) = -\log P_{\mathbf{z}}(f_{\boldsymbol{\theta}}(\mathbf{H};\mathbf{L},\mathbf{P})) - \sum_{n=0}^{N-1}\log\left|\det\frac{\partial f_{\boldsymbol{\theta}}^{n}(\mathbf{h}^n;\mathbf{L},\mathbf{P})}{\partial \mathbf{h}^n}\right|.$$

Once the optimal parameter $\theta^*$of the invertible network f_θ is learned, we can sample an HRMS from $\boldsymbol{P}_{\boldsymbol{H}|\boldsymbol{L},\boldsymbol{P}}(\boldsymbol{H} \mid \boldsymbol{L}, \boldsymbol{P};\ \boldsymbol{\theta})$, as follows:

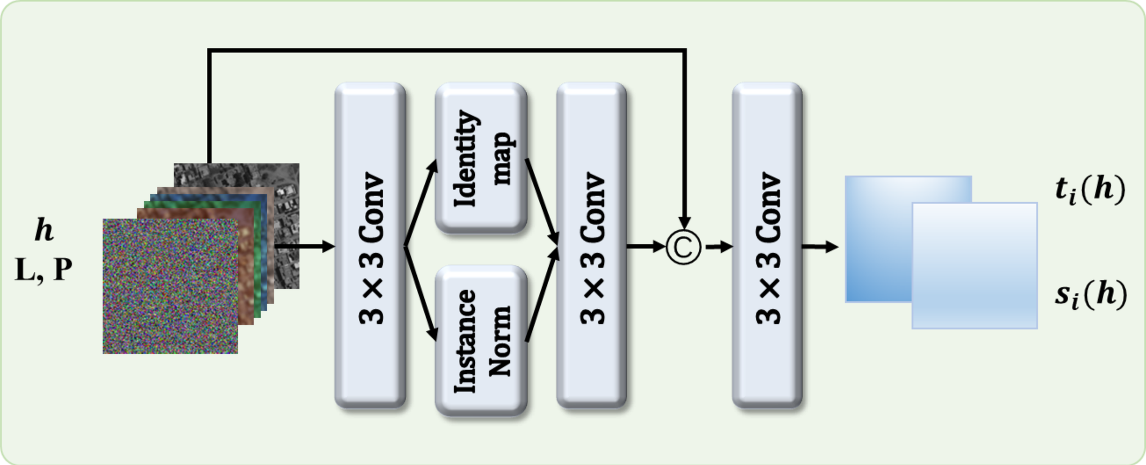
$$\mathbf{H} = f_{\boldsymbol{\theta}_*}^{-1}(\mathbf{z};\mathbf{L},\mathbf{P}),\ \ \mathbf{z} \sim \mathcal{N}(\mathbf{z}|\mathbf{0},\mathbf{I}_s).$$

Network architecture

The overall architecture



Architectures of submodules



$$\mathbf{h}_{n+1} = f_{\boldsymbol{\theta}}^{n}(\mathbf{h}_n;\mathbf{L},\mathbf{P}).$$

$$\mathbf{h}_{n+1}^1 = \mathbf{h}_n^1 \odot \exp\left(s_1\left(\mathbf{h}_n^2\right)\right) + t_1\left(\mathbf{h}_n^2\right),$$

$$\mathbf{h}_{n+1}^2 = \mathbf{h}_n^2 \odot \exp\left(s_2\left(\mathbf{h}_{n+1}^1\right)\right) + t_2\left(\mathbf{h}_{n+1}^1\right),$$

Effectiveness verification

Table 2: PSNR values of PanFlowNet with different noises.

Noise	PSNR ↑	SSIM ↑	SAM ↓	ERGAS ↓
noise 1	41.8561	0.971218	0.0223989	0.933770
noise 2	41.8581	0.971229	0.0223936	0.933516
noise 3	41.8579	0.971224	0.0223922	0.933545
noise 4	41.8563	0.971216	0.0223946	0.933642
noise 5	41.8583	0.971228	0.0223937	0.933529
noise 6	41.8552	0.971204	0.0224002	0.933823

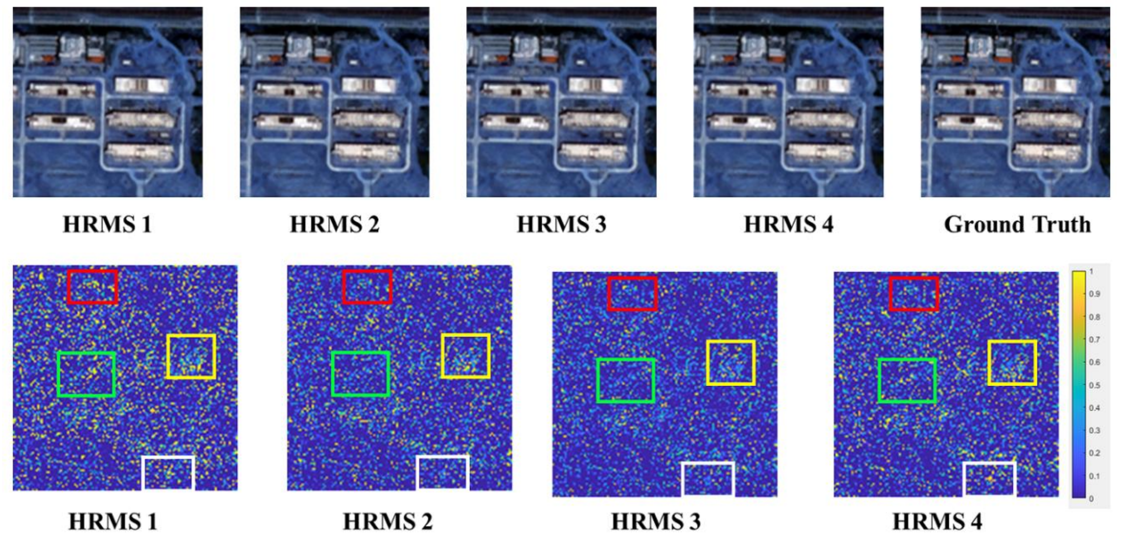


Figure 4: The visualization results are used to validate the effectiveness of our proposed PanFlowNet. The first row visualizes different HRMS images generated from different noises and gives LRMS and PAN images on the WorldView-II dataset. The second row visually shows the differences in the detailed parts that each HRMS image focuses on ground truth.

Results

Table 1: Experimental results of all the competing methods on the three benchmark datasets. The best and the second best values are highlighted in **bold** and underline, respectively.

Methods	Params	WorldView II				WorldView III				GaoFen2			
		PSNR ↑	SSIM ↑	SAM ↓	ERGAS ↓	PSNR ↑	SSIM ↑	SAM ↓	ERGAS ↓	PSNR ↑	SSIM ↑	SAM ↓	ERGAS ↓
SFIM	-	34.1297	0.8975	0.0439	2.3449	21.8212	0.5457	0.1208	8.9730	36.9060	0.8882	0.0318	1.7398
Brovey	-	35.8646	0.9216	0.0403	1.8238	22.5060	0.5466	0.1159	8.2331	37.7974	0.9026	0.0218	1.3720
GS	-	35.6376	0.9176	0.0423	1.8774	22.5608	0.5470	0.1217	8.2433	37.2260	0.9034	0.0309	1.6736
IHS	-	32.1601	**0.9812**	0.0461	2.0278	22.5579	0.5354	0.1266	8.3616	38.1754	0.9100	0.0243	1.5336
GFPCA	-	34.5581	0.9038	0.0488	2.1411	22.3344	0.4826	0.1294	8.3964	37.9443	0.9204	0.0314	1.5604
PNN	0.0689	40.7550	0.9624	0.0259	1.0646	29.9418	0.9121	0.0824	3.3206	43.1208	0.9704	0.0172	0.8528
PANNET	0.0688	40.8176	0.9626	0.0257	1.0557	29.6840	0.9072	0.0851	3.4263	43.0659	0.9685	0.0178	0.8577
MSDCNN	0.2390	41.3355	0.9664	0.0242	0.9940	30.3038	0.9184	0.0782	3.1884	45.6874	0.9827	0.0135	0.6389
SRPPNN	1.7114	41.4538	0.9679	0.0233	0.9899	30.4346	0.9202	0.0770	3.1553	47.1998	0.9877	0.0106	0.5586
GPPNN	0.1198	41.1622	0.9684	0.0244	1.0315	30.1785	0.9175	0.0776	3.2593	44.2145	0.9815	0.0137	0.7361
Ours	0.0873	**41.8584**	0.9712	**0.0224**	**0.9335**	**30.4873**	**0.9221**	**0.0751**	**3.1142**	**47.2533**	**0.9884**	**0.0103**	**0.5512**

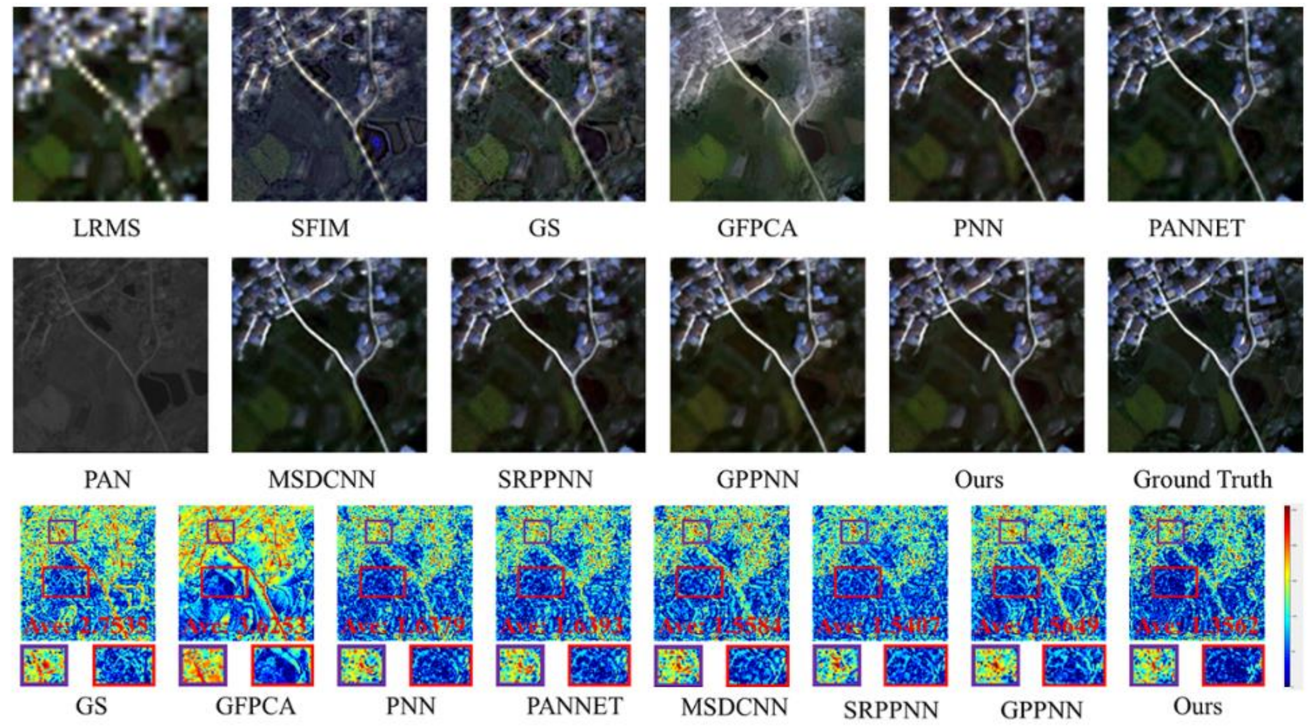


Figure 5: Visual comparison of all the competing methods on WorldViewII. The last row visualizes the error maps and average errors between the pan-sharpening results and the ground truth.

Ablations

Table 4: PSNR values of PanFlowNet with different number of stages on WorldViewII. The best and the second best values are highlighted in **bold** and underline, respectively.

Stages (K)	PSNR ↑	SSIM ↑	SAM ↓	ERGAS ↓
1	38.2469	0.9471	0.0344	1.4294
2	40.7152	0.9639	0.0255	0.9935
3	41.2664	0.9674	0.0236	0.9935
4	**41.8584**	**0.9712**	**0.0224**	**0.9335**

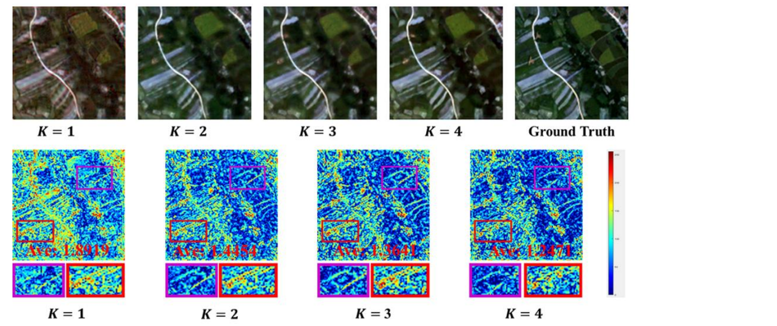


Figure 6: Intermediate visual results of different numbers of CACB in our PanFlowNet on WorldViewII. The last row visualizes the error maps and average errors between the pan-sharpening results and the ground truth.

Table 5: The results of different configurations on WorldViewII. The best and the second best values are highlighted in **bold** and underline, respectively. (PS: Parameters Sharing)

Configuration	L	P	PS	PSNR ↑	SSIM ↑	SAM ↓	ERGAS ↓
I	✗	✗	✓	31.3136	0.9033	0.0840	3.2813
II	✓	✗	✓	36.1760	0.9058	0.0315	1.6287
III	✗	✓	✓	40.8503	0.9647	0.0253	1.0539
IV	✓	✓	✗	**42.0865**	**0.9719**	**0.0215**	**0.9062**
PanFlowNet(Ours)	✓	✓	✓	41.8584	0.9712	0.0224	0.9335

code