2. Methods

In this section, we introduce the proposed dual-generator translation network fusing texture and structure features for SAR–optical image matching. As illustrated in Figure 2, the dual generators provide feedback to each other to obtain the structure and texture features, which are fused by the Bi-GFF and CFA modules. In the following subsections, we present the details of the generators, discriminator, and loss functions.

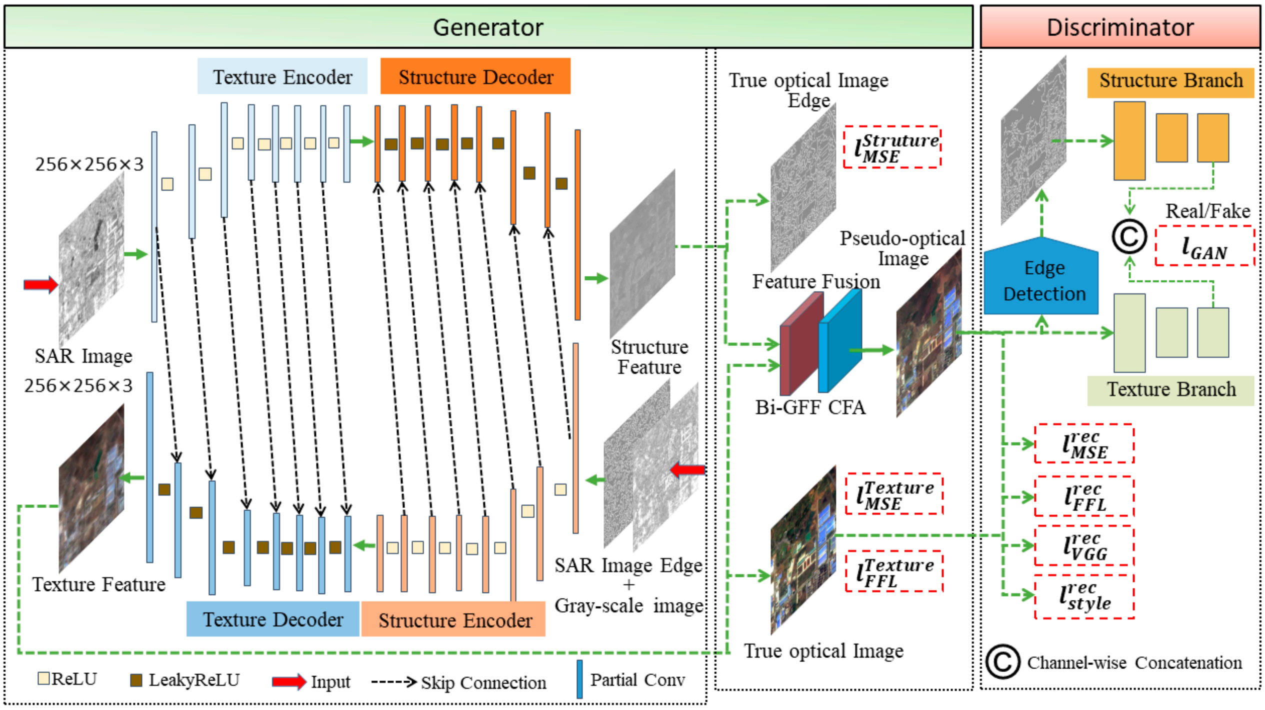

Figure 2. The generators and discriminator of our network. **Generators**: The SAR-to-optical translation process is divided between two generators, i.e., a structure generator and a texture generator, which borrow each other's depth features, and the Bi-GFF and CFA modules are used to refine and fuse the features from these structure and texture reconstruction branches to form the final pseudo-optical image. **Discriminator**: The texture branch guides texture generation, and the structure branch guides structure generation.

2.1. Generators

As shown in Figure 2, the generator part of the SAR-to-optical translation network is divided into two generators, namely, a structure generator and a texture generator, which are based on U-Net variants [42], where final features from the structure encoder and multilevel features from the texture encoder are added to the texture decoder via a skip connection, and final features from the texture encoder and multilevel features from the structure encoder are added to the structure decoder via a skip connection. We also show the structural details of the texture and structure generators in Table 1. In the encoder stage, the SAR image to be translated is passed to the texture encoder, and the greyscale image and edge structure image of the SAR image to be translated are passed to the structure encoder. In the decoder stage, the structure features from the structure encoder are used as constraints in the texture decoder, and the texture features from the texture encoder are used as constraints in the structure decoder. This coupled dual-generator structure ensures good complementarity between the structure and texture features. Compared with normal convolutional layers, partial convolutional layers can better capture the information of irregular boundaries [42]; accordingly, considering the severe scattering noise, NRD, and large irradiance differences between optical and SAR images, we also use partial convolutional layers instead of normal convolutional layers. In addition, we add skip connections in the CFA module to join together low-level and high-level features during the fusion of the structure and texture features to ensure robust prediction results.

Table 1. Details of the texture and structure generator architecture. PConv is defined as a partial convolutional layer with the specified filter size, stride, and padding. Concat indicates that structure features and texture features are connected by a skip connection.

Module Name	Filter Size	Channel	Stride	Padding	Nonlinearity
Texture/Structure (T/S) Encoder					
T/S Input		3/2			
T/S Encoder PConv1	7 × 7	64	2	3	ReLU
T/S Encoder PConv2	5 × 5	128	2	2	ReLU
T/S Encoder PConv3	5 × 5	256	2	2	ReLU
T/S Encoder PConv4	3 × 3	512	2	1	ReLU
T/S Encoder PConv5	3 × 3	512	2	1	ReLU
T/S Encoder PConv6	3 × 3	512	2	1	ReLU
T/S Encoder PConv7	3 × 3	512	2	1	ReLU
Texture Decoder					
S Encoder-PConv7		512	-	-	-
Concat (S Encoder-PConv7, T Encoder-PConv6)		512 + 512	-	-	-
T Decoder PConv8	3 × 3	512	1	1	LeakyReLU
Concat (T Decoder PConv8, T Encoder-PConv5)		512 + 512	-	-	-
T Decoder PConv9	3 × 3	512	1	1	LeakyReLU
Concat (T Decoder PConv9, T Encoder-PConv4)		512 + 512	-	-	-
T Decoder PConv10	3 × 3	512	1	1	LeakyReLU
Concat (T Decoder PConv10, T Encoder-PConv3)		512 + 256	-	-	-
T Decoder PConv11	3 × 3	256	1	1	LeakyReLU
Concat (T Decoder PConv11, T Encoder-PConv2)		256 + 128	-	-	-
T Decoder PConv12	3 × 3	128	1	1	LeakyReLU
Concat (T Decoder PConv12, T Encoder-PConv1)		128 + 64	-	-	-
T Decoder PConv13	3 × 3	64	1	1	LeakyReLU
Concat (T Decoder PConv13, T Input)		64 + 3	-	-	-
Texture Feature	3 × 3	64	1	1	LeakyReLU
Structure Decoder					
T Encoder-PConv7		512	-	-	-
Concat (T Encoder-PConv7, S Encoder-PConv6)		512 + 512	-	-	-
S Decoder PConv14	3 × 3	512	1	1	LeakyReLU
Concat (S Decoder PConv14, T Encoder-PConv5)		512 + 512	-	-	-
S Decoder PConv15	3 × 3	512	1	1	LeakyReLU
Concat (S Decoder PConv15, T Encoder-PConv4)		512 + 512	-	-	-
S Decoder PConv16	3 × 3	512	1	1	LeakyReLU
Concat (S Decoder PConv16, T Encoder-PConv3)		512 + 256	-	-	-
S Decoder PConv17	3 × 3	256	1	1	LeakyReLU
Concat (S Decoder PConv17, T Encoder-PConv2)		256 + 128	-	-	-
S Decoder PConv18	3 × 3	128	1	1	LeakyReLU
Concat (S Decoder PConv18, T Encoder-PConv1)		128 + 64	-	-	-
S Decoder PConv19	3 × 3	64	1	1	LeakyReLU
Concat (S Decoder PConv19, S Input)		64 + 2	-	-	-
Structure Feature	3 × 3	64	1	1	LeakyReLU