

# Tripartite-GAN: Synthesizing Liver Contrast-Enhanced MRI to Improve Tumor Detection

Jianfeng Zhao<sup>a</sup>, Dengwang Li<sup>a,\*</sup>, Zahra Kassam<sup>b</sup>, Joanne Howey<sup>b</sup>, Jaron Chong<sup>c</sup>, Bo Chen<sup>b,d</sup>, Shuo Li<sup>b,d,\*</sup>

<sup>a</sup>*Shandong Key Laboratory of Medical Physics and Image Processing, Shandong Institute of Industrial Technology for Health Sciences and Precision Medicine, School of Physics and Electronics, Shandong Normal University, Jinan, 250358, China*

<sup>b</sup>*Department of Medical Imaging, Western University, London, Canada*

<sup>c</sup>*McGill University, Montral QC, Canada*

<sup>d</sup>*Digital Imaging Group of London, Canada*

---

## Abstract

Contrast-enhanced magnetic resonance imaging (CEMRI) is crucial for the diagnosis of patients with liver tumors, especially for the detection of benign tumors and malignant tumors. However, it suffers from high-risk, time-consuming, and expensive in current clinical diagnosis due to the use of the gadolinium-based contrast agent (CA) injection. If the CEMRI can be synthesized without CA injection, there is no doubt that it will greatly optimize the diagnosis. In this study, we propose a Tripartite Generative Adversarial Network (Tripartite-GAN) as a non-invasive, time-saving, and inexpensive clinical tool by synthesizing CEMRI to detect tumors without CA injection. Specifically, our innovative Tripartite-GAN combines three associated-networks (an attention-aware generator, a convolutional neural network-based discriminator, and a region-based convolutional neural network-based detector) for the first time, which achieves CEMRI synthesis and tumor detection promoting each other in an end-to-end framework. The generator facilitates detector for accurate tumor detection via synthesizing tumor-specific CEMRI. The detector promotes the generator for accurate CEMRI synthesis via the back-propagation. In order to synthesize CEMRI of equivalent clinical value to real CEMRI, the attention-aware generator expands the receptive field via hybrid convolution, and enhances feature representation and context learning of multi-class liver MRI via dual attention mechanism, and improves the performance of convergence of loss via residual learning. Moreover, the attention maps obtained from the generator newly added into the detector improve the performance of tumor detection. The discriminator promotes the generator to synthesize high-quality CEMRI via the adversarial learning strategy. This framework is tested on a large corpus of axial T1 FS Pre-Contrast MRI and axial T1 FS Delay MRI of 265 subjects. Ex-

---

\*Corresponding author

Email addresses: dengwang@sdnu.edu.cn (Dengwang Li), slishuo@gmail.com (Shuo Li)

perimental results and quantitative evaluation demonstrate that the Tripartite-GAN achieves high-quality CEMRI synthesis that peak signal-to-noise rate of 28.8 and accurate tumor detection that accuracy of 89.4%, which reveals that Tripartite-GAN can aid in the clinical diagnosis of liver tumors.

**Keywords:** Contrast-enhanced MRI synthesis; Tripartite-GAN; Dual attention module; Tumor detection.

## 1. Introduction

2 The magnetic resonance examination of liver cancer relies heavily on contrast agent (CA) injection. Specifically, as shown in Fig.1, in the non-contrast enhanced magnetic resonance imaging (NCEMRI) obtained without CA injection, the area of hemangioma (a benign tumor) and hepatocellular carcinoma (HCC, a malignant tumor) could barely find their difference to distinguish. On the contrary, in contrast-enhanced MRI (CEMRI) obtained by the CA injection, the area of hemangioma is gradual central filling and bright at the edge and the area of HCC is entirely or mostly bright through the whole tumor. It is no doubt that CA injection gives the two kinds of tumors their diagnosis specificity, which provides an accurate and easy way to diagnose hemangioma and HCC.

13 However, gadolinium-based CA brings inevitable shortcomings, which suffers from high-risk, time-consuming, and expensive (Idée et al., 2006). The high-risk is due to the gadolinium-based CA injection, which may induce nephrogenic systemic fibrosis (Marckmann et al., 2006), especially for patients with compromised kidney function. The time-consuming comes from the MRI process itself and the waiting-time after CA injection. The expensive mainly comes from CA, in the USA alone, conservatively, if each dose of CA is \$60, the direct material expense alone equates to roughly \$1.2 billion in 2016 (Statistics from IQ-AI Limited Company, USA). It will be significant for clinical diagnosis if CEMRI can be successfully synthesized without CA injection. Therefore, this work focuses on the end-to-end method of synthesizing liver CEMRI from liver NCEMRI for tumor detection.

25 There is currently no reported synthesis of liver CEMRI for tumor detection because of three unique challenges: 1) The difficulty in discriminating the tumor features extracted in NCEMRI. That is to say. It is easy to confuse the features of hemangioma and HCC when extracting the features because of the low discrimination of hemangioma and HCC in NCEMRI. Therefore, the synthesis network is required to pay more attention to the detailed feature for improving the feature representation. 2) The difficulty in learning the highly nonlinear mapping between multi-class NCEMRI and multi-class CEMRI. That is to say. Each anatomy can be seen as a class. Different from the synthesis of single-class medical images (e.g., brain MRI (Yang et al., 2018), lesion area patches (Frid-Adar et al., 2018)), liver MRI has multi-class anatomy (i.e., liver, spleen, spine, and so on). Therefore, multi-class liver MRI has the risk of causing

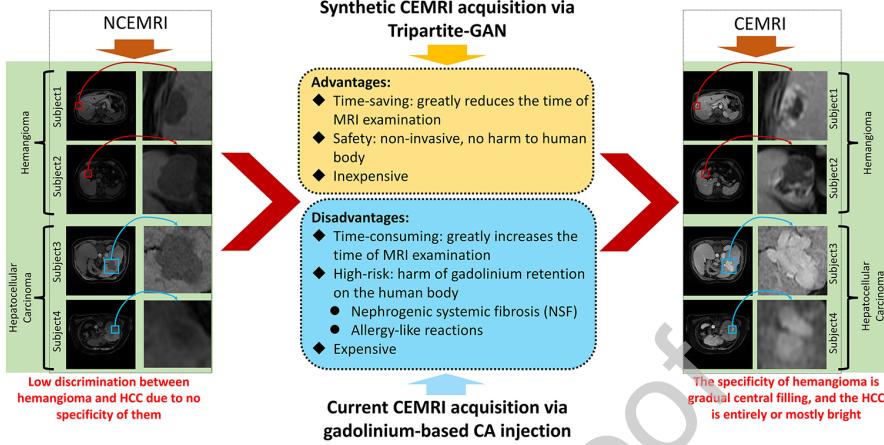


Figure 1: Our method of obtaining CEMRI without CA injection has advantages of non-invasive, time-saving, and inexpensive. There are four cases of synthesizing CEMRI from NCEMRI. Subject1 and Subject2 are hemangioma, a benign tumor. Subject3 and Subject4 are hepatocellular carcinoma(HCC), a malignant tumor.

misclassification of the anatomy. The synthesis network is required to explicitly capture global dependencies of multi-class feature representations regardless of locations. 3) The difficulty in alleviating the blurring problem of synthetic CEMRI. The blurring of synthetic image is a problem that GAN always needs to alleviate (Korkinof et al., 2018), and the problem becomes more serious due to the complex anatomy of multi-class liver MRI. That is to say. For our task of synthesizing liver CEMRI to improve tumor detection, we must ensure not only the high quality of CEMRI synthesis but also the clarity of the tumor area. Therefore, an effective loss function is necessary for our task due to the loss function determines the ability to learn the highly nonlinear mapping between the source image and the target image.

In this paper, we propose a novel Tripartite Generative Adversarial Network (Tripartite-GAN) as a non-invasive, time-saving, and inexpensive clinical tool to synthesize liver CEMRI without CA injection for tumor detection. Specifically, for the first time, the Tripartite-GAN combines three associated-network (an attention-aware generator, a convolutional neural network-based (CNN-based) discriminator, and a region-based convolutional neural network-based (R-CNN-based) detector), which simultaneously achieves CEMRI synthesis and tumor detection in an end-to-end framework. Firstly, in order to overcome the aforementioned challenges of 1) and 2), the newly designed attention-aware generator expands the receptive field via hybrid convolution, integrates local features with their global dependencies via dual attention module (DAM), and improves the convergence of loss via residual learning. This is capable of effectively extracting the diagnosis-specific features of two types of tumor and accurately learning the highly nonlinear mapping between multi-class NCEMRI and multi-class

62 CEMRI. Secondly, in order to overcome the aforementioned challenge of 3) for  
 63 achieving high-quality CEMRI synthesis, which equivalent to real CEMRI. The  
 64 CNN-based discriminator is trained to discriminate the real or fake of synthetic  
 65 CEMRI, and then promotes the generator to synthesize highly authentic CEMRI  
 66 via adversarial-strategy. Thirdly, the R-CNN-based detector is combined to the  
 67 generator via back-propagation for the first time, which achieves that CEMRI  
 68 synthesis and tumor detection promote each other in an end-to-end framework.  
 69 Moreover, the attention maps obtained from the generator newly added into the  
 70 detector improve the performance of tumor detection. The contributions of this  
 71 study are mainly in four aspects:

- 72 1. For the first time, synthesizing CEMRI without CA injection for liver  
 73 tumor detection is achieved, which provides a safe, time-saving, and inex-  
 74 pensive clinical tool to synthesize CEMRI without CA injection.
- 75 2. The newly proposed Tripartite-GAN successfully combined the regular  
 76 two-participant GAN and the detector via back-propagation for the first  
 77 time, which achieves that CEMRI synthesis and tumor detection promote  
 78 each other in an end-to-end framework.
- 79 3. The newly designed attention-aware generator is powerful in feature ex-  
 80 traction with the help of hybrid convolution, residual learning, and DAM.  
 81 Specifically, the hybrid convolution enlarges the receptive field efficiently,  
 82 the residual learning benefits the convergence to facilitate the training of  
 83 the generator, and the DAM enhances feature representation learning of  
 84 tumor specificity and context learning of multi-class liver MRI.
- 85 4. Attention maps from the generator newly added into the detector in the  
 86 manner of residual connection improve VGG-16 based convolution opera-  
 87 tion to extract tumor information better, which improves the performance  
 88 of tumor detection.

## 89 **2. Related work**

### 90 *2.1. Related work on tumor diagnosis in liver MRI*

91 Studies have shown that in liver tumor diagnosis, MRI is useful and more  
 92 sensitive than other modalities in the tumor diagnosis and characterization of  
 93 hemangioma and HCC with the help of CA injection (Digumarthy et al., 2005;  
 94 Low, 2007; Halavaara et al., 2006). However, the dependence of existing mag-  
 95 netic resonance examination on CA injection also brings inevitable shortcoming,  
 96 which suffers from high-risk, time-consuming, and expensive (Idée et al., 2006).  
 97 Therefore, there is an urgent need for a technique to synthesize image with the  
 98 same clinical effect as real CEMRI.

### 99 *2.2. Related work on automated analysis in liver MRI*

100 To the best of our knowledge, although there are many works focus on medi-  
 101 cal image synthesis (Yang et al., 2018; Costa et al., 2017), no work has attempted  
 102 on liver CEMRI synthesis from NCEMRI without CA injection due to unique

103 challenges comes from complex anatomy and patient diversity in liver MRI.  
 104 And no work has attempted that CEMRI synthesis and tumor detection pro-  
 105 mote each other in an end-to-end framework. Existing works include but limited  
 106 to medical image synthesis and tumor detection are done separately, while the  
 107 synthesis work is mostly focused on the single-class medical image (e.g., brain  
 108 MRI (Yang et al., 2018), lesion area patch (Frid-Adar et al., 2018)).

109 *2.2.1. Traditional GAN for medical image analysis*

110 GAN has demonstrated great power in the medical image analysis since it  
 111 was proposed by Goodfellow et al.(Goodfellow et al., 2014), which is used to  
 112 model the image distribution of generated samples to be indistinguishable from  
 113 target images. There are many studies focusing on medical image synthesis  
 114 have obtained certain success (Nie et al., 2018; Beers et al., 2018; Emami et al.,  
 115 2018). Besides, based on the generated samples, a wide variety of applications  
 116 are derived. Such as improving liver lesion classification via GAN-based syn-  
 117 synthetic image (Frid-Adar et al., 2018), improving the accuracy and clarity of  
 118 retinal vessel segmentation by using GAN-based network (Son et al., 2017), ac-  
 119 celerating automatic spondylolisthesis grading from MRIs across modalities by  
 120 using a customized GAN, (Zhao et al., 2019) and improving lesion detection by  
 121 using GAN-based network to synthesise PET from CT (Ben-Cohen et al., 2019).  
 122 These GAN-based works highlight the importance of image synthesis quality.  
 123 It is worth noting that these works focus more on high-quality medical image  
 124 synthesis, and then using the generated samples to improve the associated tasks.

125 Although these works attempt to use GAN to promote another associated  
 126 network, the GAN and the associated network work separately. None of them  
 127 achieve the combination and mutual improvement between GAN and other as-  
 128 sociated networks in an end-to-end framework. Recently, in the field of natural  
 129 images, some studies have attempted to combine GAN and other tasks networks  
 130 and obtained some success (Simon et al., 2019; Shen et al., 2018; Chongxuan  
 131 et al., 2017). For instance, in (Simon et al., 2019), they proposed a three-player  
 132 GAN to improve classification networks, which combined the GAN and classi-  
 133 fier by back-propagation. Inspired by this, in this paper, we propose a novel  
 134 Tripartite-GAN for that liver CEMRI synthesis and tumor detection promote  
 135 each other in an end-to-end framework.

136 *2.2.2. Attention module in networks*

137 Since the work (Vaswani et al., 2017) proposed to use the self-attention  
 138 mechanism to draw global dependencies of inputs, which successfully applied  
 139 to the machine translation, the attention mechanism has been widely used in  
 140 various deep learning-based tasks. For example, the work (Zhang et al., 2018)  
 141 proposed a self-attention GAN to model long-range dependencies effectively. In  
 142 (Hu et al., 2018a), they proposed relation modules to learn the information be-  
 143 tween objects for improving object recognition in an end-to-end object detector.  
 144 In particular, the work (Fu et al., 2019) embedded channel attention module and  
 145 position attention module into the network of scene segmentation for adaptively  
 146 integrate local features with their global dependencies. In our Tripartite-GAN,

147 the DAM is applied successfully to the GAN-based medical image synthesis task  
 148 for the first time, which motivated by the success of attention modules in the  
 149 above works.

150 **3. Methodology**

151 *3.1. An overview of Tripartite-GAN*

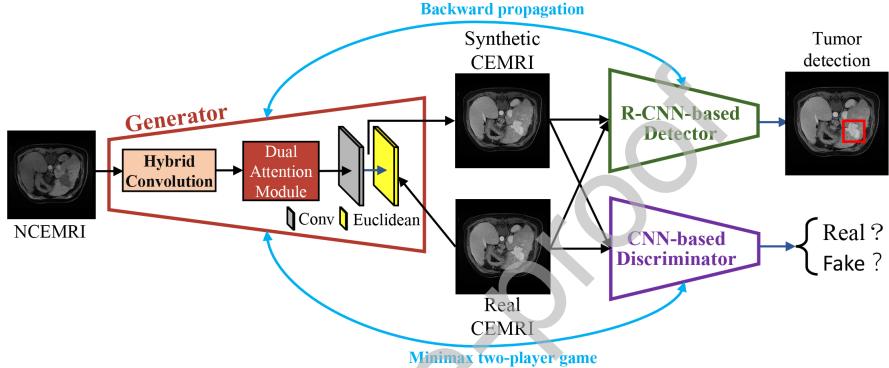


Figure 2: The Tripartite-GAN synthesizes CEMRI for tumor detection by the combination of three associated-task networks, the attention-aware generator, the CNN-based discriminator and the R-CNN-based detector. The R-CNN-based detector directly detects tumor from the synthetic CEMRI and improves the accuracy of synthetic CEMRI generation via back-propagation. The CNN-based discriminator urges generator to generate more realistic synthetic CEMRI via the adversarial-strategy.

152 For effective CEMRI synthesis and tumor detection, our Tripartite-GAN ex-  
 153 ecutes the competition between three participants: the newly designed attention-  
 154 aware generator (section 3.2), the CNN-based discriminator(section 3.3), and  
 155 the R-CNN-based detector (section 3.4). Fig.2 shows the design of our newly  
 156 proposed Tripartite-GAN. Specifically, the attention-aware generator is a hy-  
 157 brid convolution network to synthesize tumor-specific CEMRI, which facilitates  
 158 tumor detection. The generator is embedded the DAM which consists of a  
 159 minutious attention module (MAM, section 3.2.2) and a global attention mod-  
 160 ule (GAM, section 3.2.3) in a parallel manner. The MAM exploits the interde-  
 161 pendencies between channel feature maps to emphasize interdependent feature  
 162 maps and improve the feature representation of specific anatomy, which can  
 163 enhance the detailed synthesis of tumor specificity. The GAM encodes a wider  
 164 range of contextual information of multi-class liver MRI into local features to  
 165 make the Tripartite-GAN context-aware, which enhances the learning ability  
 166 of highly nonlinear mapping between multi-class NCEMRI and CEMRI, And  
 167 then improves the accuracy and spatial continuity of CEMRI synthesis. Be-  
 168 sides, the hybrid convolution was adopted to reserve the useful information  
 169 without the pooling operation and expanding the receptive field in the step of  
 170 feature extraction. The discriminator is a CNN designed to distinguish real or

171 fake of synthetic CEMRI and promote the generator to synthesize high-quality  
 172 CEMRI via adversarial-strategy. The detector is a R-CNN designed to locate  
 173 and classify the tumor. The detector prompts the generator to focus on the  
 174 difference between NCEMRI and CEMRI via the back-propagation, especially  
 175 for the specificity of hemangioma and HCC. Meanwhile, the generator improves  
 176 the performance of tumor detection via newly adding attention maps into the  
 177 detector in the manner of residual connection.

178 *3.2. Attention-aware generator with DAM for CEMRI synthesis*

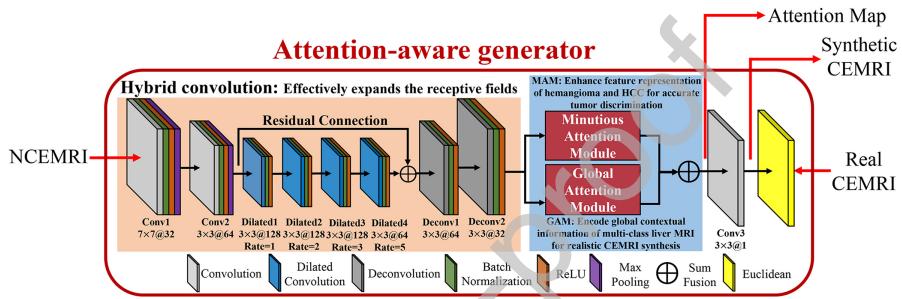


Figure 3: The generator aims to synthesis accurate and realistic CEMRI. It uses a hybrid convolution consists of standard convolution layers, dilated convolution layers, and deconvolution layers. The dilated convolution is utilized to enlarge receptive fields. The two standard convolution layers and two deconvolution layers are connected to the front and back of dilated convolution, which reduces the size of feature maps to expand the receptive fields more efficiently. Followed the hybrid convolution, the DAM (MAM and GAM) enhances the detailed feature extraction and aggregates long-range contextual information of the generator, which improves the detailed synthesis of the specificity of the tumor and the spatial continuity of the multi-class liver MRI.

179 The attention-aware generator is designed based on a fully convolutional net-  
 180 work (FCN) by embedding DAM in a parallel manner, which aims to synthesize  
 181 CEMRI from NCEMRI by learning the nonlinear mapping between CEMRI and  
 182 NCEMRI. Firstly, the generator extracts feature from NCEMRI via hybrid con-  
 183 volution. Moreover, the application of residual learning in CNN has achieved  
 184 promising results in many challenging generic image processing tasks (He et al.,  
 185 2016). For the generator in our Tripartite-GAN, the residual learning is per-  
 186 formed to connect the layer of Conv2 and the layer of Dilated4, which improves  
 187 the convergence of generator loss to facilitate the training of the generator. Sec-  
 188 ondly, the feature map obtained from the layer of Deconv2 is fed into the DAM  
 189 (MAM and GAM), the MAM enhances detailed feature extraction by utilizing  
 190 the interdependencies between channel maps of the layer of Deconv2, and the  
 191 GAM explicitly captures global dependencies of multi-class liver MRI by en-  
 192 coding global contextual information into local features. Followed the DAM,  
 193 we perform an element-wise sum to accomplish the feature fusion of MAM and  
 194 GAM. Lastly, the last layer of Conv3 is used to generate the final synthetic  
 195 CEMRI.

196 *3.2.1. The architecture of the attention-aware generator*

197 As shown in Fig.3, the generator synthesizes CEMRI mainly goes through  
 198 three steps: Hybrid convolution, DAM, and the last convolution layer. Specifi-  
 199 cally, the generator adopts four dilated convolution layers rather than standard  
 200 convolution layers for enlarging receptive fields. Besides, two convolution lay-  
 201 ers and two deconvolution layers are added to the front end and back end of  
 202 dilated convolution operations, respectively. The two convolution layers con-  
 203 tain the operation of convolution, batch normalization (BN), Rectified Linear  
 204 Unit (ReLU), and max-pooling. The four dilated convolution layers contain the  
 205 operation of dilated convolution, BN, and ReLU. The two deconvolution layers  
 206 contain the operation of deconvolution, BN, and ReLU. Note that we do not  
 207 adopt all the eight layers as dilated convolution. This reduces the size of feature  
 208 maps and increases the receptive field more efficiently. The kernel sizes from  
 209 the first convolution operation Conv1 to the last convolution operation Conv3  
 210 are  $7 \times 7$ ,  $3 \times 3$  and  $3 \times 3$ , respectively.  
 211 The numbers of filters are 32, 64, 128, 128, 128, 64, 64, 32 and 1, respectively.  
 212 Besides, inspired by the hybrid dilated convolution (Wang et al., 2018), the di-  
 213 lated rate are set to 1, 2, 3, and 5 from the layer of Dilated1 to Dilated4 for  
 214 avoiding the gridding issue.

215 *3.2.2. MAM: Enhancing the feature representation of hemangioma and HCC  
 216 for accurate tumor discrimination*

217 The MAM explicitly models the interdependencies between channels of the  
 218 Deconv2 in hybrid convolution. For the hybrid convolution, each channel map  
 219 of high-level features can be regarded as an anatomy-specific response, and the  
 220 different anatomic structure responses are associated with each other (Hu et al.,  
 221 2018b; Fu et al., 2019). Therefore, the MAM emphasizes interdependent feature  
 222 maps and improves the feature representation of specific anatomy by utilizing  
 223 the interdependencies between channel maps. Especially for the CEMRI, the  
 224 difference between the specificity of tumors and normal tissues are more conspic-  
 225 uous than the NCEMRI. Accordingly, the MAM is embedded into the generator  
 226 to enhance detailed feature extraction, especially for the details of tumor speci-  
 227 ficity. Specifically, after the feature map of the Deconv2 feeding into MAM, the  
 228 MAM goes through three steps to obtain the output feature. Firstly, a channel  
 229 attention matrix is generated, which models the channel relationship between  
 230 any two pixels of the feature map. Secondly, a matrix multiplication operation  
 231 is performed between the channel matrix and the original features. Thirdly,  
 232 an element-wise sum operation is performed on the above multiplied resulting  
 233 matrix and original features to obtain the final representations reflecting the  
 234 specificity of the different anatomy.

235 The MAM as shown in Fig.4, we directly calculate the minutious attention  
 236 feature map  $Z \in \mathbb{R}^{C \times C}$  from the feature  $X \in \mathbb{R}^{C \times H \times W}$  of Deconv2. Firstly, we  
 237 reshape  $X$  to  $\mathbb{R}^{C \times N}$ , and then perform a matrix multiplication between  $X$  and  
 238 its transpose. Next, we apply a softmax layer to obtain the minutious attention

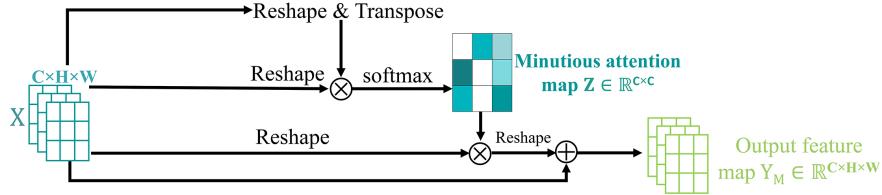


Figure 4: The details of MAM, it enhances the detailed feature extraction by utilizing the interdependencies between channel maps  $X$ .

239 map  $Z \in \mathbb{R}^{C \times C}$ :

$$\text{Minutious Attention Matrix : } Z_{ji} = \frac{\exp(X_i \cdot X_j)}{\sum_{i=1}^C \exp(X_i \cdot X_j)} \quad (1)$$

240 where  $Z_{ji}$  measures the impact of  $i^{th}$  channel on  $j^{th}$  channel. In addition, we  
241 perform a matrix multiplication between the transpose of  $Z$  and  $X$ , then reshape  
242 their result to  $\mathbb{R}^{C \times H \times W}$ . Lastly, we multiply the result by a scale parameter  $\beta$   
243 and use an execution element summation operation to get the final output  $Y_M$ :

$$\text{Output of MAM : } Y_{Mj} = \beta \sum_{i=1}^C (Z_{ji} X_i) + X_j \quad (2)$$

244 where  $\beta$  is initialized as 0 and gradually increase weight through learning. The  
245  $Y_M$  shows the final feature of each channel is a weighted sum of the features of  
246 all channels and original features, it boosts feature discriminability.

### 247 3.2.3. GAM: Aggregating long-range contextual information of multi-class liver 248 MRI for CEMRI synthesis

249 Context relationship is essential for anatomic structure understanding, es-  
250 pecially for the liver MRI with complex anatomical structures. However, many  
251 works (Peng et al., 2017; Zhao et al., 2017) suggest that traditional FCN could  
252 lead to misclassification of objects with local feature representations generated.  
253 To overcome the defect of local feature representations, the GAM explicitly  
254 captures global dependencies regardless of locations, which adaptively aggre-  
255 gate long-range contextual information to make the framework context-aware.  
256 Specifically, after the feature map of the Deconv2 feeding into GAM, the GAM  
257 goes through three steps to obtain the output feature, which is similar to the  
258 MAM. The first step is to generate a spatial attention matrix that models the  
259 spatial relationship between any two pixels of the feature map. Secondly, a  
260 matrix multiplication operation is performed between the spatial matrix and  
261 the original features. Thirdly, an element-wise sum operation is performed on  
262 the above multiplied resulting matrix and original features to obtain the final  
263 representations reflecting long-range contexts.

264 The GAM encodes the global contextual information into local features, thus  
265 enhancing their representative capability. The operation of GAM as illustrated

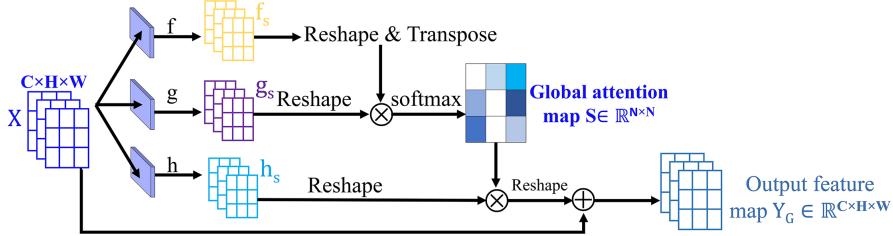


Figure 5: The details of GAM, it explicitly captures global dependencies of multi-class liver MRI by encoding global contextual information into local features.

in Fig.5, first give the feature  $X \in \mathbb{R}^{C \times H \times W}$  of Deconv2 to the GAM. Then feed  $X$  into two parallel convolution layers  $f$  and  $g$  to generate two feature map  $f_s$  and  $g_s$ , where the kernel size of layer  $f$  and layer  $g$  are all  $1 \times 1$ . Next, we reshape  $f_s$  and  $g_s$  from  $\mathbb{R}^{C \times H \times W}$  to  $\mathbb{R}^{C \times N}$ , where  $N = H \times W$  is the number of features. After that we perform a matrix multiplication between the transpose of  $f_s$  and  $g_s$ , and apply a softmax layer to calculate the global attention map  $S \in \mathbb{R}^{N \times N}$ :

$$\text{Global Attention Matrix : } S_{mn} = \frac{\exp(f_{sn} \cdot g_{sm})}{\sum_{i=1}^N \exp(f_{sn} \cdot g_{sm})} \quad (3)$$

where  $S_{mn}$  measures the impact of  $n^{th}$  position on  $m^{th}$  position. Note that the more similar feature representations of the two position contributes to greater correlation between them.

The bottom road is a convolution layer with  $1 \times 1$  kernel size. It is used to generate a feature map  $h_s \in \mathbb{R}^{C \times H \times W}$ , and then reshape it to  $h_s \in \mathbb{R}^{C \times N}$ . Next, we perform a matrix multiplication between  $h_s$  and the transpose of  $S$ , and then reshape the result to  $\mathbb{R}^{C \times H \times W}$ . Finally, we multiply it by a scale parameter  $\alpha$  and perform a element-wise sum operation with the features  $X$  to obtain the final output feature maps  $Y_G \in \mathbb{R}^{C \times H \times W}$  as follows:

$$\text{Output of GAM : } Y_{Gm} = \alpha \sum_{n=1}^N (S_{mn} h_{sn}) + X_m \quad (4)$$

where  $\alpha$  is initialized as 0 and gradually increase weight through learning. In equation 4, the final feature  $Y_G$  at each position is a weighted sum of the features at all positions and the original feature  $X$ . Therefore, it has a global contextual view and selectively aggregates contexts according to the global attention map. These feature representations achieve mutual gains and more robust for the CEMRI synthesis.

The primary advantages of the attention-aware generator include:

- Hybrid convolution expands the receptive fields more efficiently.
- DAM enhances the ability to extract features by modeling the interdependencies between channels and encoding global contextual information into

292 local features.

- 293 • Residual learning facilitates the convergence of the training loss of our  
294 Tripartite-GAN, which makes the loss of Tripartite-GAN lower.

295 *3.3. The CNN-based discriminator makes the Tripartite-GAN adversarial-strategy-  
296 aware*

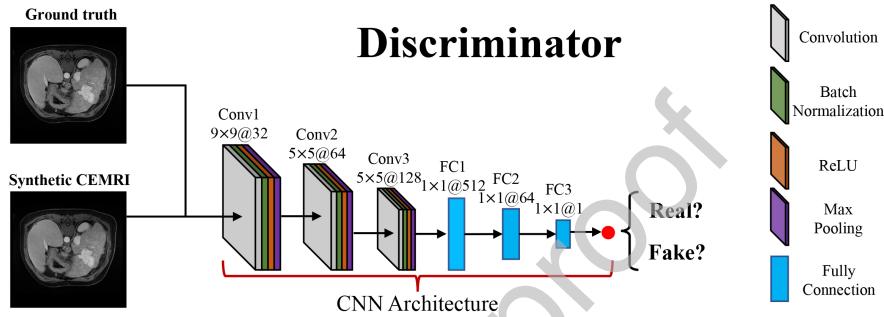


Figure 6: The CNN architecture of the discriminative network receives the ground truth of CEMRI and the synthetic CEMRI, and then outputs the discriminative results of real or fake. Its adversarial strategy eagerly supervises attention-aware generator to find its own mistakes, which increased the authenticity of the synthetic CEMRI.

297 As shown in Fig.6, the CNN-based discriminator consists of three convolutional  
298 layers and three fully connected layers, where each convolution is followed  
299 by the operation of BN, ReLU, and max-pooling. The discriminator makes the  
300 Tripartite-GAN adversarial-strategy-aware because it receives either the syn-  
301 thetic CEMRI from the generator or real CEMRI from ground truth, and out-  
302 puts a single scalar to indicate the image is real or fake and feedbacks to the  
303 generator. With learning through this confrontation, the discriminative eagerly  
304 prompts the generator to synthesize highly realistic CEMRI, until the truth or  
305 false is hard to discriminate.

306 *3.4. The detector is the first time combined with the regular GAN in an end-to-  
307 end framework for tumor detection*

308 Fig.7 shows the architecture of the detector, which is a customized Faster  
309 R-CNN (Ren et al., 2015). The detector aims to directly locate the tumor  
310 location for obtaining the ROI of hemangioma and HCC. And the detector dis-  
311 tinguishes whether it is a benign tumor or a malignant tumor. Meanwhile, in  
312 this Tripartite-GAN, the detector encourages generator to pay more attention to  
313 the specificity of the two types of tumors via back-propagation, which makes the  
314 generator to synthesize CEMRI accurately. The detection is mainly comprised  
315 of two steps. The first stage is proposing the candidate tumor bounding boxes  
316 via the Region Proposal Network. The second stage is performing classification

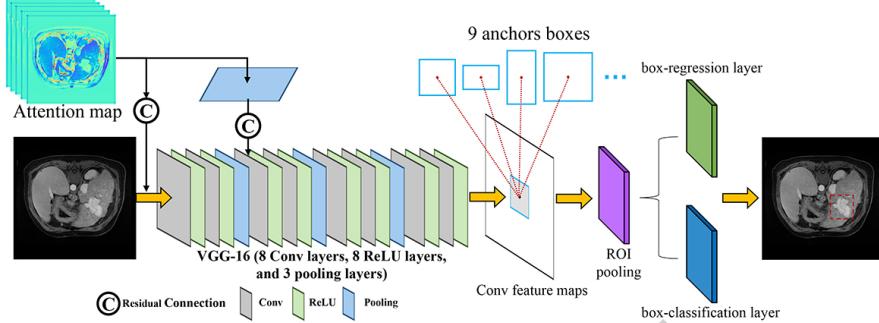


Figure 7: The architecture of the tumor detection network, which receives synthetic CEMRI and then accurately localizes the Region of Interest (RoI) of the tumor and classifies the tumor. Attention maps from the generator newly added into the detector in the manner of residual connection improve VGG-16 based convolution operation to extract tumor information better, which improves the performance of tumor detection. Meanwhile, the back-propagation of  $L_{cls}$  prompts the generator to focus on the specificity between two types of tumors. Added  $L_{cls}$  into Tripartite-GAN achieves a win-win between detector and generator via back-propagation.

and bounding-box regression on the RoI of the tumor proposed in the first stage. It adopts a customized Simonyan and Zisserman model (Simonyan and Zisserman, 2014) (VGG-16) with the help of attention maps obtained from DAM in the generator for feature extraction, in which the VGG-16 based network has eight shareable convolutional layers.

The advantage over a single detection network is that we can get help from the attention maps extracted by the attention-aware generator. Although the tasks of the generator and detector are different, the generator is used to synthesize CEMRI, and the detector is used to detect tumors. While the generator extracts the global features of the entire liver MRI to ensure the quality of the entire synthetic CEMRI, it can also extract local features (such as tumor-specific features). Similar to the generator, the detector needs to extract the tumor feature for tumor location and classification. Therefore, the features extracted by the generator can help the feature extraction of the detector in the characterization of tumor details. In our Tripartite-GAN, we achieve the help of the generator to the detector by adding the attention maps to VGG-16 based convolution operation in the manner of residual connection. It mainly has two shortcuts of adding the attention map to VGG-16. The first shortcut is to add the attention map from DAM directly before the first convolution layer of VGG-16, and the second shortcut is to add the attention map after the pooling layer operation to the third convolution layer of VGG-16.

### 3.5. Tripartite loss joint strategy combines three associated-tasks

In order to synthesize high-quality CEMRI of equivalent clinical value to real CEMRI, and then promote tumor detection. Tripartite-GAN uses a tripartite loss function  $L_G$  to train the attention-aware generator. The tripartite loss mainly consists of three items corresponding to three robust losses of three

343 tasks. The three tasks are the synthesis of CEMRI, discrimination of CEMRI,  
 344 and classification of hemangioma and HCC. It means that the generator not only  
 345 mutually promote with discriminator by adversarial strategy but also mutually  
 346 optimizes with the detector by back-propagation. The tripartite loss  $L_G$  of  
 347 the generator is shown in equation5. The first item is an Euclidean loss  $L_E$ ,  
 348 which used to maximize peak signal-to-noise rate (PSNR) for producing high-  
 349 quality synthetic-CEMRI. The second item is Cross-Entropy loss  $L_{CE}$ . Like the  
 350 learning strategy of traditional GAN, we perform the two-participant minimax  
 351 game between generator and discriminator by using loss  $L_{CE}$ . The loss  $L_{CE}$   
 352 is used to minimize the probability of the samples generated by the generator  
 353 to be recognized while maximizing the probability of the discriminator making  
 354 a mistake (Mahasseni et al., 2017). In other words, through the adversarial  
 355 learning strategy, the ability of the generator to synthesize CEMRI, and the  
 356 ability of discriminator to discriminate real or fake are improved simultaneously.  
 357 The third item is the loss function of  $L_{cls}$  for training detector. It helps to  
 358 optimize the softmax-based tumor classification. The  $L_G$  is a weighted sum of  
 359 three items comprising  $L_E$ ,  $L_{CE}$  and  $L_{cls}$ . The tripartite loss adopted to train  
 360 the generator has a stable performance, which is formulated as follow:

$$L_G(X, Y, p, u) = \underbrace{L_E(X, Y)}_{\text{Synthesis loss}} + \lambda_1 \underbrace{L_{CE}(D(G(X), 1))}_{\text{Discrimination loss}} + \lambda_2 \underbrace{L_{cls}(p, u)}_{\text{Classification loss}} \quad (5)$$

361 where the hyper-parameter  $\lambda_1$  and  $\lambda_2$  are used for maintaining the weight of  
 362 joint learning of adversarial learning and back-propagation of  $L_{cls}$ . The  $G(X)$   
 363 is the synthetic CEMRI from NCEMRI ( $X$ ) by the generator. And  $Y$  repre-  
 364 sents the real CEMRI, which is the ground truth. The  $D(G(X))$  is the prob-  
 365 ability computed by the discriminator, and the value of  $D(G(X))$  was taken  
 366 into 0 or 1 (0 corresponds to fake, and 1 corresponds to real). Meanwhile,  
 367 the  $L_{CE}(D(G(X), 1))$  function promotes generator to produce more realistic  
 368 CEMRI for confusing the discriminator, and it makes the network adversarial-  
 369 strategy-wise. The loss function  $L_E$  and  $L_{CE}$  as follows:

$$L_E(X, Y) = \|Y - G(X)\|_2^2 \quad (6)$$

$$L_{CE}(\hat{Y}, Y) = - \sum_i Y_i \log(\hat{Y}_i) + (1 - Y_i) \log(1 - \hat{Y}_i) \quad (7)$$

370 where  $Y$  is the real CEMRI and  $\hat{Y}$  is the synthetic-CEMRI by the generator,  
 371 and the loss function of  $L_D$  for training discriminator is defined as:

$$L_D(X, Y) = L_{CE}(D(Y), 1) + L_{CE}(D(G(X)), 0) \quad (8)$$

372 that is, the principle of the discriminator is similar to a classifier, one classifier  
 373 classifies the  $X$  as ‘real’ or ‘fake’. The third item  $L_{cls}$  of tripartite loss is one

374 part of the detection loss  $L_{D_e}$ , which is a multi-task loss to jointly trained for  
 375 tumor classification and bounding-box regression. The multi-task loss  $L_{D_e}$  can  
 376 be defined as follow:

$$L_{D_e}(p, u, t^u, v) = \overbrace{L_{cls}(p, u)}^{\text{Classification loss}} + \lambda_3 \overbrace{[u \geq 1] L_{box}(t^u, v)}^{\text{Bounding-box loss}} \quad (9)$$

377 where the hyper-parameter  $\lambda_3$  set to one for maintaining the balance of two  
 378 tasks losses of  $L_{cls}$  and  $L_{box}$ . The classification loss and bounding-box loss  $L_{box}$   
 379 are identical as those defined in Fast R-CNN (Girshick, 2015):

$$\begin{cases} L_{cls}(p, u) = -\log p_u \\ L_{box}(t^u, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L_1}(t_i^u - v_i) \end{cases} \quad (10)$$

380 in which

$$smooth_{L_1}(x) = \begin{cases} 0.5x^2 & if |x| < 1 \\ |x| - 0.5 & otherwise, \end{cases} \quad (11)$$

381 where the  $p$  represents the probability distribution of RoI of the tumor,  $u$  rep-  
 382 resents which type of tumor belongs to, the  $[u \geq 1]$  evaluates to 1 when  $u \geq 1$   
 383 and 0 otherwise.  $t^u$  is the predicted tuple of bounding-box, and  $v$  is a true tuple  
 384 of the bounding-box.

385 The primary advantages of tripartite loss not only have a stable performance  
 386 but also successfully achieves that liver CEMRI synthesis, CEMRI discrimina-  
 387 tion, and tumor detection promote each other in an end-to-end framework.

## 388 4. Experiments and Results

### 389 4.1. Materials and Implementation

390 The experimental datasets we used totaling 265 subjects (75 subjects of  
 391 hemangioma, 138 subjects of HCC, and 52 subjects of health), and each sub-  
 392 ject has corresponding NCEMRI and CEMRI (after gadolinium CA injection)  
 393 collected after standard clinical liver MRI examinations. And all subjects are  
 394 provided after approval by the McGill University Health Centre. The corre-  
 395 sponding axial T1 FS Pre-Contrast MRI [4mm; 512×512px] and axial T1 FS  
 396 Delay MRI [4mm; 512×512px] are selected for our experiments, in which axial  
 397 T1 FS Pre-Contrast MRI is used as NCEMRI and axial T1 FS Delay MRI is  
 398 used as CEMRI.

399 Specifically, we perform one 5-fold cross-validation test to train our Tripartite  
 400 GAN for performance evaluation and comparison. The 265 subjects are divided  
 401 into 5 groups following random rules grouping, and each group contains 53  
 402 subjects. Each of the first four groups contains 15 subjects of hemangioma,  
 403 28 subjects of HCC, and 10 subjects of health. And the last group contains  
 404 15 subjects of hemangioma, 26 subjects of HCC, and 12 subjects of health.

---

**Algorithm 1** Tripartite generative adversarial network

---

**Input:** A dataset of non-contrast MRI  $x$ ; Ground truth maps of contrast-enhanced MRI  $y$ ; The label of the types of tumor  $u$ ; The true tuple of the bounding-box  $v$ ; The loss balanced weights  $\lambda_1, \lambda_2, \lambda_3$ ; Learning rates  $\eta_1, \eta_2, \eta_3$ ; Mini-batch size  $n$ ; The number of iterations  $M$ ;

**output:** Learned parameters  $\{\theta_g, \theta_{di}, \theta_{de}, p, t^u\}$ ;

- 1: Initialize the parameters  $\{\theta_g, \theta_{di}, \theta_{de}\}$  randomly and construct model graph;
- 2: **for** step in  $M$  **do**
- 3:   fed  $u, v, x_n, y_n \leftarrow x_n$  and  $y_n$  represent the  $x$  and  $y$  with mini-batch size  $n$ ;
- 4:   /\* The **forward** propagation of  $G(x_n)$ : \*/
- 5:    $X_n = Hybrid-Conv(x_n)$ ;
- 6:    $Atten_n = MAM(X_n) + GAM(X_n)$ ;
- 7:    $G(x_n) = Conv3(Atten_n)$ ;
- 8:   /\* The **forward** propagation of  $D(\cdot)$ : \*/
- 9:    $D(G(x_n)) = CNN(G(x_n))$ ;
- 10:    $D(y_n) = CNN(y_n)$ ;
- 11:   /\* The **forward** propagation of  $D_e(\cdot)$ : \*/
- 12:    $p, t^u = R-CNN(G(x_n) + Atten_n)$
- 13:   /\* The **backward** propagation of  $G(x_n)$ : \*/
- 14:    $\theta_g = \theta_g - \eta_1 \nabla (L_E(x_n, y_n) + \lambda_1 L_{CE}(D(G(x_n)), 1) + \lambda_2 L_{cls}(p, u))$
- 15:   /\* The **backward** propagation of  $D(\cdot)$ : \*/
- 16:    $\theta_{di} = \theta_{di} - \eta_2 \nabla (L_{CE}(D(y_n), 1) + L_{CE}(D(G(x_n)), 0))$ ;
- 17:   /\* The **backward** propagation of  $D_e(\cdot)$ : \*/
- 18:    $\theta_{de} = \theta_{de} - \eta_3 \nabla L_{D_e}(p, u, t^u, v)$ ;
- 19: **end for**

---

405 In our experiments, 4 groups were used for training and 1 group was used for  
 406 testing. Then executed this process 5 times in a loop, until each group is used  
 407 as the training and testing object. Inspired of (Simon et al., 2019), the values  
 408 of hyper-parameter  $\lambda_1$  in equation 5 is set to one, and  $\lambda_2$  in equation 5 updated  
 409 according to scheme (Springenberg, 2015):

$$\lambda_2 = 1 - \frac{2\omega}{1 + \exp(-10t)} \quad (12)$$

410 where  $t$  is reducing linearly from 1 to 0 during training progress, the value of  $\omega$  is  
 411 set to 0.5 smaller than one to raise the priority of high-quality CEMRI synthesis.  
 412 The  $\lambda_2$  gradually grows during training, ensuring the weight of classification  
 413 increases with the quality of the synthetic-CEMRI. Inspired by (Girshick, 2015),  
 414 the hyper-parameter  $\lambda_3$  in  $L_{D_e}$  (9) is set to one in all experiments for maintaining  
 415 the balance two tasks of bounding-box regression and tumor classification. The  
 416 Tripartite-GAN is implemented on Pytorch library by using a server platform  
 417 with four Tesla P100 GPUs.

#### 418 4.2. Results evaluation and comparison

##### 419 4.2.1. Accurate CEMRI synthesis

420 Results of synthetic-CEMRI obtained by Tripartite-GAN are shown in Fig.8.  
 421 The synthetic-CEMRI has the equivalent value of real CEMRI in clinical diag-  
 422 nosis, in which the differences of specificity between hemangioma and HCC in  
 423 synthetic-CEMRI are accurate and obvious. The area of hemangioma is grad-  
 424 ual central filling and bright at the edge, and the area of HCC is entirely or  
 425 mostly bright through the whole tumor. The results prove that the synthetic-  
 426 CEMRI has an equal diagnostic value to the real CEMRI (ground truth) via  
 427 CA injection visually.

428 To quantitatively evaluate the synthetic performance of our Tripartite-GAN,  
 429 the Tripartite-GAN is compared with three synthesis methods: atlas-based  
 430 method (Vercauteren et al., 2009), conditional generative adversarial nets (CGANs)  
 431 (Mirza and Osindero, 2014) and Auto-Context based GAN (AC-GAN) (Nie  
 432 et al., 2018). The results of the comparison are shown in Fig.9. The pixel inten-  
 433 sity curve and zoomed local patches of tumor area show that our Tripartite-GAN  
 434 is more accurate than three other methods.

435 Besides, we also perform the ablation studies to prove every part of the  
 436 newly designed Tripartite-GAN contributes to CEMRI synthesis. The ablation  
 437 studies include Tripartite-GAN without discriminator (No  $D$ ), without DAM  
 438 (No DAM), without MAM (No MAM), without GAM (No GAM), without  
 439 detector (No  $D_e$ ), without dilated convolution (No Di-con), and without residual  
 440 learning (No Res-L). The results of ablation studies are shown in Fig.10. The  
 441 pixel intensity curve and zoomed local patches of tumor areas show that every  
 442 part of the newly designed Tripartite-GAN improves CEMRI synthesis.

443 The standard metric of normalized mean absolute error (MAE) and PSNR as  
 444 the standard of evaluation are used to evaluate the synthetic-CEMRI (Wolterink  
 445 et al., 2017), which as shown in Table1. The mean MAE and the mean PSNR

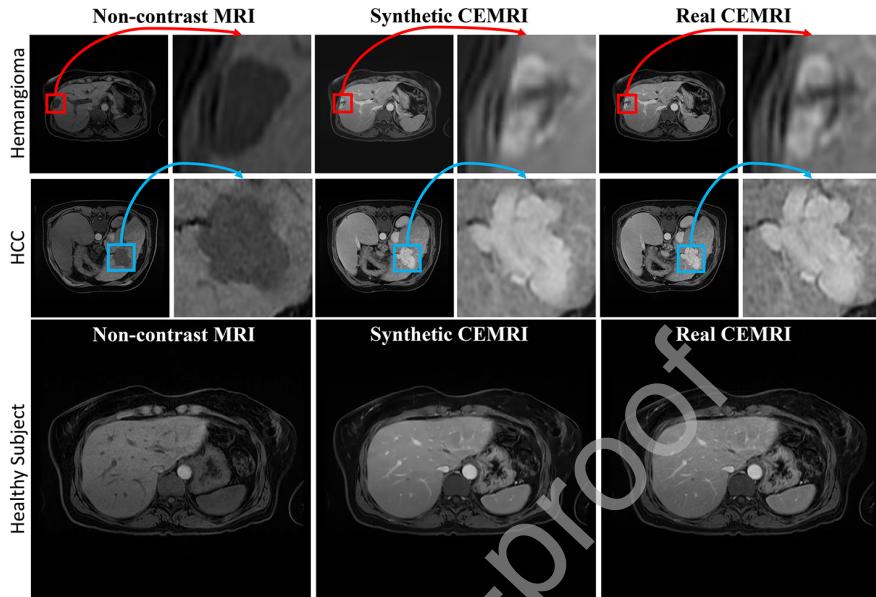


Figure 8: Our method synthesized CEMRI has an equal diagnostic value to real CEMRI. In the first two-row, it is clear that the area of hemangioma becomes gradual central filling and bright at the edge in synthetic CEMRI, and the area of HCC becomes entirely or mostly bright through the whole tumor. The red and blue windows represent the hemangioma and HCC, respectively, and enlarge them on the right. The last row is the synthesis result of healthy subjects.

of our method achieve 125.8 and 28.8, which is the best among three synthesis methods. And the evaluation of ablation studies demonstrate that every part of the newly designed Tripartite-GAN improves CEMRI synthesis.

Moreover, as shown in Table2, the paired t-tests between our proposed method and the other three works were performed on both PSNR and MAE values. The  $p$ -Values ( $<0.05$ ) of paired t-tests show that the difference between our proposed method and the other three methods is significant.

#### 4.2.2. DAM enhances feature representation: GAM improves the spatial continuity and MAM enhances the detailed synthesis

In order to verify the contribution of the DAM to CEMRI synthesis, we perform the comparison of Tripartite-GAN without DAM and our proposed Tripartite-GAN. When DAM is removed, the PSNR value decreases from 28.8 to 26.1. As the first row shown in Fig.11, the subject1 demonstrate that Tripartite-GAN with DAM outperforms Tripartite-GAN without DAM in the detailed synthesis of anatomy specificity and the spatial continuity. To verify the respective contributions of GAM and MAM to CEMRI synthesis, we perform two comparisons of evaluating the contribution of MAM and GAM independently. One of the two comparisons is between Tripartite-GAN without GAM and our

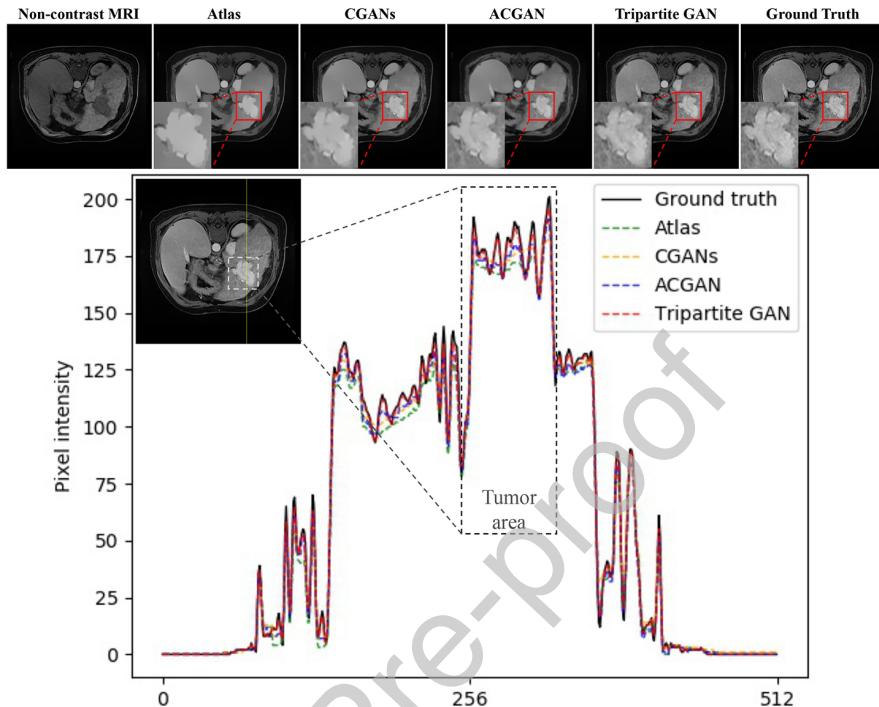


Figure 9: In the aspect of the detailed expression of the tumor and the highly realistic synthetic-CEMRI, Tripartite-GAN outperforms three other methods. The pixel intensity curve and zoomed local patches of tumor area show that our Tripartite-GAN is more accurate than three other methods.

464 Tripartite-GAN. The other is between Tripartite-GAN without MAM and our  
 465 Tripartite-GAN. When MAM is removed, the PSNR value decreases from 28.8  
 466 to 26.5. And when GAM is removed, the PSNR value decreases from 28.8 to  
 467 26.7. As the last two rows are shown in Fig.11, the subject2 demonstrates  
 468 that GAM improves the spatial continuity of CEMRI synthesis. The subject3  
 469 demonstrates that MAM enhances the detailed feature extraction to improve  
 470 the discrimination of hemangioma and HCC. The subject3 shows the failure of  
 471 not being able to differentiate HCC and Hemangioma when MAM is removed,  
 472 which incorrectly synthesizes the specificity of hemangioma into the specificity  
 473 of HCC. It is obvious that the results prove that the MAM enhances the detailed  
 474 synthesis of anatomy specificity, and the GAM improves the spatial continuity  
 475 effectively.

476 In order to clearly see the specific contributions of MAM and GAM, Fig.12  
 477 shows the feature maps of Tripartite-GAN without GAM, Tripartite-GAN with-  
 478 out MAM, and our Tripartite-GAN (with DAM), respectively. We can see the  
 479 GAM enhances the spatial feature extraction (especially for the feature extrac-  
 480 tion at edges) by comparing the red window in the feature maps. And we can

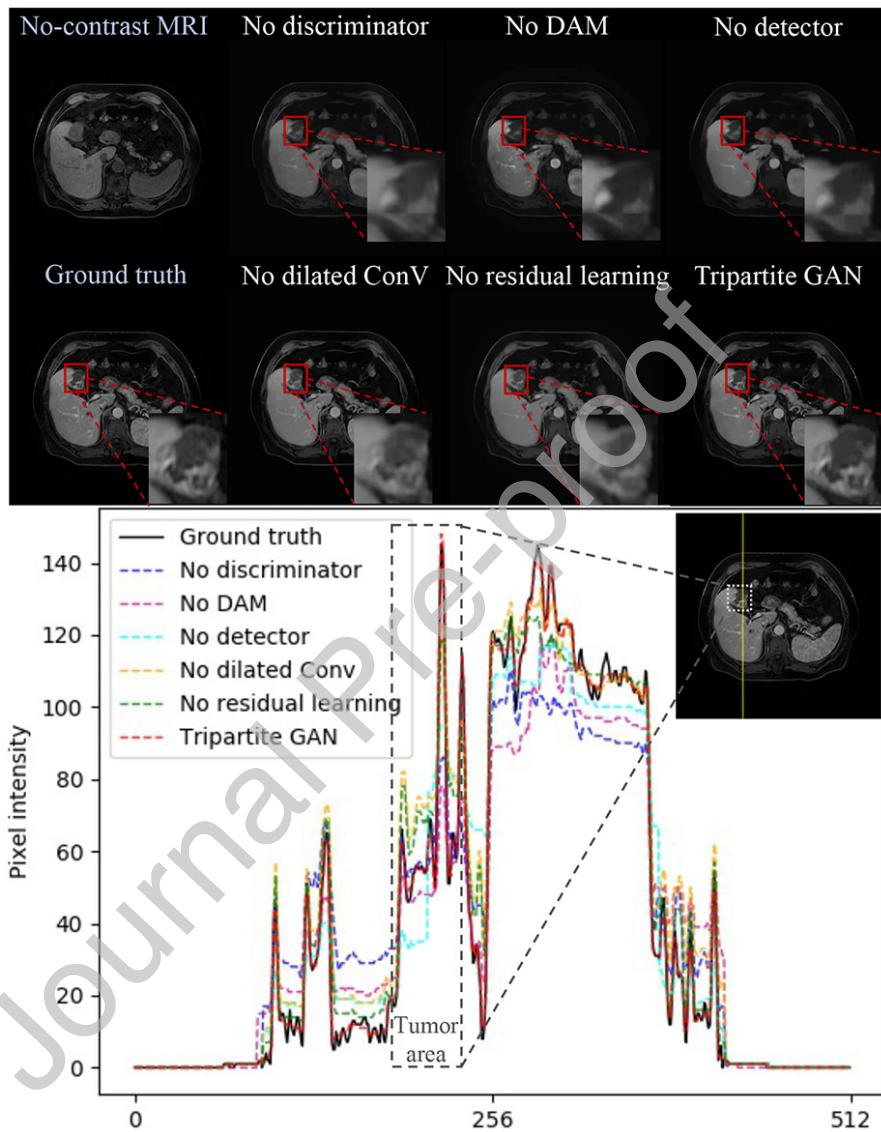


Figure 10: The ablation studies of No discriminator, No DAM, No detector, No dilated convolution, and No residual learning, which demonstrate that every part of the newly designed Tripartite-GAN contributes to CEMRI synthesis. The pixel intensity curve and zoomed local patches of tumor area demonstrate that our proposed Tripartite-GAN is more accurate and more powerful in the detailed synthesis. The horizontal coordinate denotes pixel positions of the yellow line drawn in the ground truth, and the vertical coordinate is the pixel intensity of the corresponding pixel.

Table 1: The comparison of Tripartite-GAN and three other methods of image-to-image translation demonstrates that our proposed Tripartite-GAN outperforms the three others on average MAE and PSNR. The ablation studies of No  $D$ , No DAM, No MAM, No GAM, No  $D_e$ , No Di-con, and No Res-L demonstrate that our proposed Tripartite-GAN every part of the newly designed Tripartite-GAN improves CEMRI synthesis.

	MAE		PSNR	
	Mean(std)	Med	Mean(std)	Med
<i>Comparison</i>				
Atlas	190.6(39.2)	191.2	21.8(1.6)	21.7
CGANs	177.4(34.8)	177.8	23.3(2.4)	23.2
ACGAN	162.1(28.6)	162.3	25.5(2.2)	25.2
<i>Ablation study</i>				
No $D$	167.9(25.3)	168.6	25.9(2.1)	25.5
No DAM	159.3(22.1)	159.1	26.1(1.9)	26.0
No MAM	157.4(20.4)	157.0	26.5(1.7)	26.7
No GAM	155.2(20.0)	155.5	26.7(1.7)	26.8
No $D_e$	152.3(18.9)	151.9	26.9(1.6)	26.6
No Di-con	145.7(18.2)	148.1	27.3(1.5)	27.4
No Res-L	139.7(17.3)	140.0	27.8(1.4)	27.6
<b>Tripartite-GAN</b>	<b>125.8(16.2)</b>	<b>125.3</b>	<b>28.8(1.4)</b>	<b>28.8</b>

Table 2:  $p$ -Values by performing paired t-tests between our Tripartite-GAN and other related works for both PSNR and MAE values.

	MAE	PSNR
Atlas	<0.01	<0.01
CGANs	<0.01	<0.01
ACGAN	<0.05	<0.05

481 see the MAM enhances the detailed feature extraction by comparing the green  
 482 window in the feature maps. It is clear that GAM helps our Tripartite-GAN  
 483 adaptively aggregates long-range contextual information, which improves the  
 484 spatial continuity of synthetic CEMRI. And MAM helps our Tripartite-GAN  
 485 enhances detailed feature extraction, which ensures the accurate synthesis of  
 486 the specificity of tumor in synthetic CEMRI. The visualization of synthesis re-  
 487 sults and zoomed local patches of tumor area are shown in the last two columns  
 488 in Fig.12. It is clear that MAM helps our Tripartite-GAN enhance detailed  
 489 synthesis, and GAM helps our Tripartite-GAN improve the spatial continuity  
 490 of synthetic CEMRI.

#### 491 4.2.3. Hybrid convolution for increasing the effective receptive field

492 In order to verify the advantages of our designed hybrid convolution for  
 493 CEMRI synthesis, we perform the comparison between a traditional FCN (No

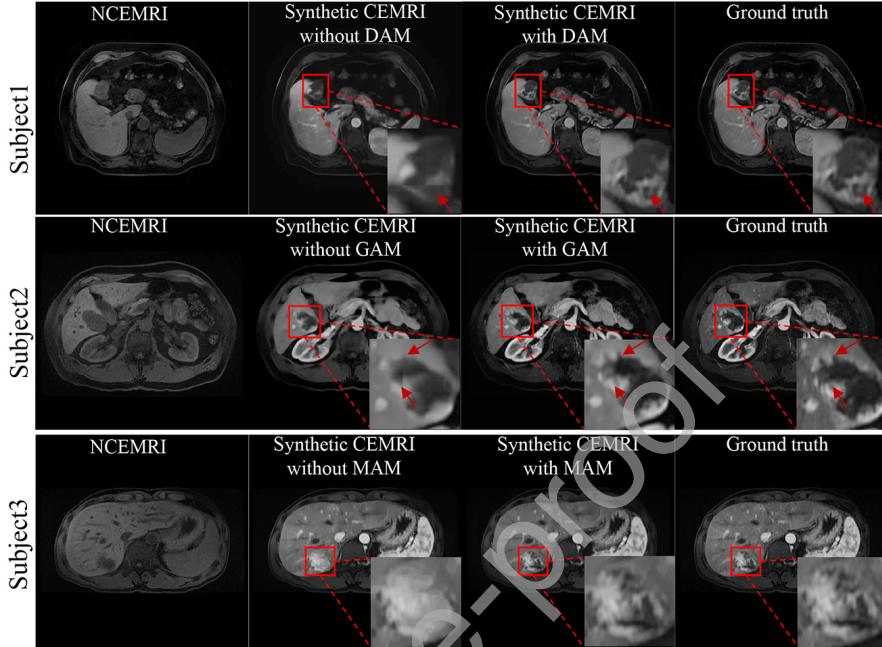


Figure 11: The subject1 demonstrates that DAM enhances the detailed synthesis of anatomy specificity and the spatial continuity. The subject2 demonstrates that GAM improves the spatial continuity of CEMRI synthesis. The subject3 demonstrates that MAM enhances the detailed feature extraction to improve the discrimination of hemangioma and HCC. The subject3 shows the failure of not being able to differentiate HCC and Hemangioma when MAM is removed, which incorrectly synthesizes the specificity of hemangioma into the specificity of HCC. The red windows of zoomed local patches represent the tumor area. From left to right, they are the NCEMRI, the synthetic CEMRI without attention module, the synthetic CEMRI with attention module, and the ground truth, respectively.

494 Di-con) with the same parameter settings of dilated convolution layers and our  
 495 proposed Tripartite-GAN. The synthesis results are shown in Fig.10. And the ef-  
 496 fect of the hybrid convolution operation is quantitatively evaluated. The PSNR  
 497 value of synthetic CEMRI by using No Di-con Tripartite-GAN is 27.3, and the  
 498 PSNR value of synthetic CEMRI by using proposed Tripartite-GAN is 28.8.  
 499 These results prove that expansion of the receptive field by hybrid convolution  
 500 makes the generator powerful in feature extraction.

#### 501 4.2.4. Residual learning benefits the training of generator

502 In order to verify the effect of the residual learning to CEMRI synthesis,  
 503 we perform the comparison between Tripartite-GAN without residual learning  
 504 and our proposed Tripartite-GAN. The synthesis results are shown in Fig.10.  
 505 And the PSNR value of synthetic CEMRI from Tripartite-GAN without resi-  
 506 dential learning is 27.8, while our proposed Tripartite-GAN is 28.8, respectively.  
 507 Moreover, the visualization of the training loss is shown in Fig.13, in which the  
 508 Tripartite-GAN without residual learning and our proposed Tripartite-GAN

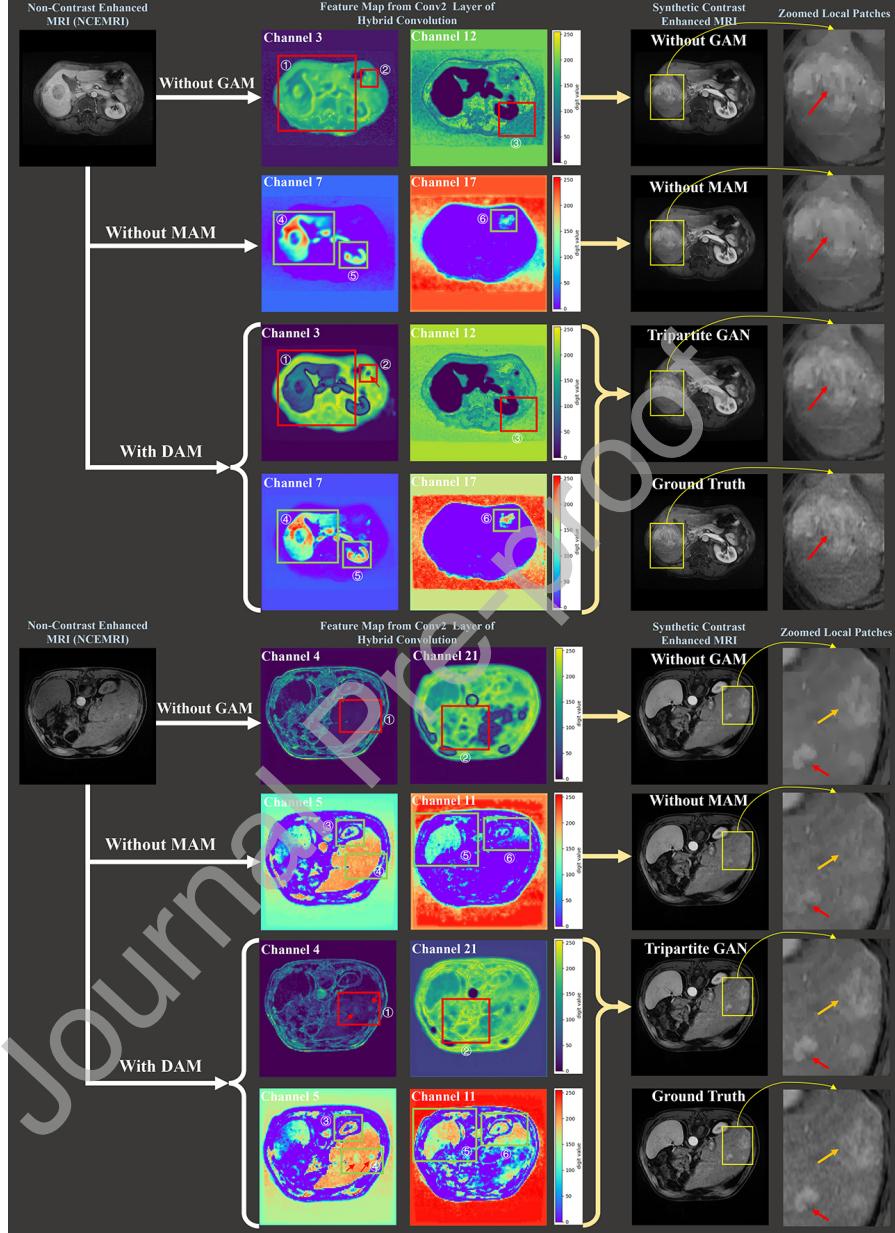


Figure 12: Two examples of CEMRI synthesis. The red windows in the feature maps represent the difference of spatial continuity between Tripartite-GAN without GAM and our Tripartite-GAN. The green windows in feature maps represent the difference of detailed feature extraction between Tripartite-GAN without MAM and our Tripartite-GAN. The last two columns show the synthesis results and zoomed local patches of the tumor area. It is clear that MAM helps our Tripartite-GAN enhance detailed synthesis, and GAM helps our Tripartite-GAN improve the spatial continuity of synthetic CEMRI.

509 are demonstrated by the green curve and red curve, respectively. It is clear  
 510 that Tripartite-GAN outperforms the Tripartite-GAN without residual learn-  
 511 ing. And the results prove that residual learning helps the network train faster  
 512 and achieve lower synthesis error.

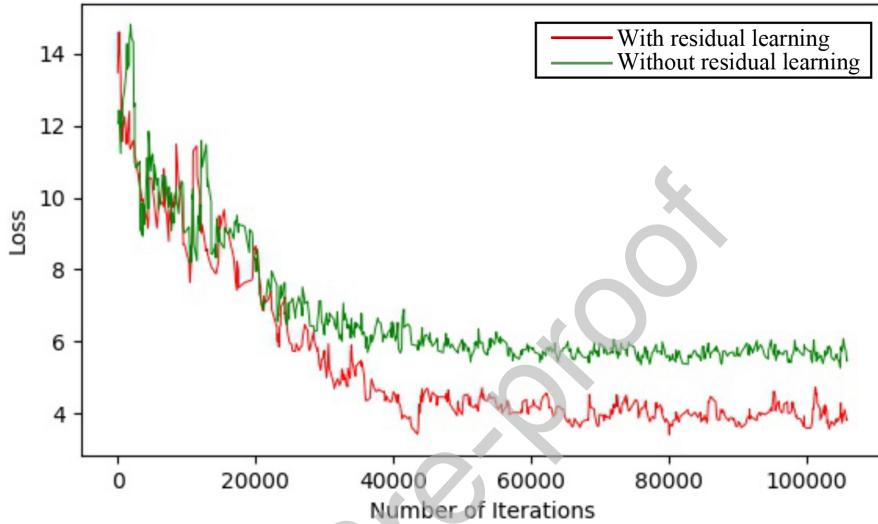


Figure 13: Our proposed newly designed generator of Tripartite-GAN has lower training loss compared with the Tripartite-GAN without residual learning.

#### 513 4.2.5. Adversarial strategy encourages the high-quality CEMRI synthesis

514 In order to verify the contribution of the adversarial strategy to CEMRI syn-  
 515 thesis, we performed a comparison between Tripartite-GAN without discrimi-  
 516 nator and proposed Tripartite-GAN, which are shown in Fig.10. Apparently,  
 517 the synthetic CEMRI of proposed Tripartite-GAN has fewer artifacts than the  
 518 Tripartite-GAN without the adversarial strategy. The quantitative evaluation  
 519 shows that the PSNR value decreases from 28.8 to 25.9 when discriminator  
 520 is removed. And these results prove that the discriminator makes the frame-  
 521 work of Tripartite-GAN adversarial-strategy-aware, which eagerly improves the  
 522 authenticity of synthetic-CEMRI of the generator.

#### 523 4.2.6. Back-propagation of classification loss urges the more accurate CEMRI 524 synthesis

525 In order to verify the contribution of the detector to the generator, we per-  
 526 form the comparison between Tripartite-GAN without the detector and our  
 527 proposed Tripartite-GAN, which are shown in Fig.10. We can clearly see that  
 528 the proposed Tripartite-GAN outperforms the Tripartite-GAN without the de-  
 529 tector. Especially for the specificity learning of the tumor. The PSNR value  
 530 decreases from 28.8 to 26.9 when the detector is removed. These results prove

531 that the detector is reinforced the performance of the attention-aware generator  
 532 via the back-propagation of  $L_{cls}$ . The detector prompts the generator to focus  
 533 on the specificity of two types of tumors. Meanwhile, the synthetic CEMRI gen-  
 534 erated by the attention-aware generator has distinct specific, which facilitates  
 535 detector for accurate detection of tumors.

536 As shown in Fig.10, we perform the a comparison by using the ablation study,  
 537 which verifies the contribution of the discriminator to the CEMRI synthesis,  
 538 the contribution of the DAM to the CEMRI synthesis, the contribution of the  
 539 detector to the CEMRI synthesis, the contribution of dilated convolution to  
 540 the CEMRI synthesis, and the contribution of residual learning to the CEMRI  
 541 synthesis. All of the pixel intensity curves and zoomed local patches of tumor  
 542 area proved that these modules of Tripartite GAN have a positive effect on  
 543 liver CEMRI synthesis, especially in improving the accuracy of the synthesis of  
 544 tumor areas.

#### 545 4.2.7. Accuracy tumor detection

546 Results of tumor detection via detector from Tripartite-GAN show that our  
 547 proposed Tripartite-GAN has a high and stable accuracy of 89.4%. To quan-  
 548 titatively evaluate the performance of detection of our Tripartite-GAN, the  
 549 Tripartite-GAN was compared with three detection methods: U-Net based FCN  
 550 (Dong et al., 2017), modified Faster-R-CNN (Akselrod-Ballin et al., 2016) and  
 551 combination of fuzzy c-means and SVM(FZM-SVM) (Singh et al., 2015). Re-  
 552 sults of the comparison of Tripartite-GAN and the other three methods of detec-  
 553 tion are shown in Table3, which demonstrates that our proposed Tripartite-GAN  
 554 outperforms the three other methods of detection. Besides, we also perform the  
 555 ablation studies to prove every part of the newly designed Tripartite-GAN con-  
 556 tributes to tumor detection. The results of ablation studies are shown in Table4.  
 557 The ablation studies include Tripartite-GAN without generator and discrimina-  
 558 tor (No  $G + No D$ ), without discriminator (No  $D$ ), without DAM (No DAM),  
 559 without MAM (No MAM), without GAM (No GAM), without adding attention  
 560 maps into detector (No atten), without dilated convolution (No Di-con), and  
 561 without residual learning (No Res-L). Results of ablation studies demonstrate  
 562 that every part of the newly designed Tripartite-GAN improves tumor detection.

Table 3: The comparison of Tripartite-GAN and three other methods of detection, which demonstrates that our proposed Tripartite-GAN outperforms three other detection methods. The upper bound of Tripartite-GAN demonstrates that our synthetic CEMRI images performed close to the real CEMRI in tumor detection.

<i>Method</i>	FZM- SVM	Modified Faster-RCNN	U-Net based FCN	<b>Tripartite GAN</b>	<b>Tripartite-GAN (upper bound)</b>
Accuracy	78.1%	79.2%	79.2%	<b>89.4%</b>	<b>90.9%</b>

563 We also evaluated the accuracy, sensitivity, and specificity of tumor detec-  
 564 tion results in Table5, Table6, and Table7. We defined healthy subjects as  
 565 positive and hemangioma as negative in Table5. We defined healthy subjects

Table 4: The ablation studies demonstrate that every part of the newly designed Tripartite-GAN contributes to tumor detection. The upper bound of Tripartite-GAN demonstrates that our synthetic CEMRI images performed close to the real CEMRI in tumor detection.

<i>Method</i>	No G + D	No D	No DAM	No MAM	No GAM	No Atten	No Di-Con	No Res-L	<b>Tripartite GAN</b>	<b>Tripartite-GAN (upper bound)</b>
Accuracy	80.0%	81.1%	83.0%	83.8%	84.9%	86.4%	87.5%	88.3%	<b>89.4%</b>	<b>90.9%</b>

566 as positive and HCC as negative in Table6. And we defined hemangioma as  
 567 positive and HCC as negative in Table7. The evaluation results demonstrate  
 568 that our Tripartite-GAN outperforms three other detection methods. And the  
 569 ablation study of Tripartite-GAN demonstrates that every part of the newly  
 570 designed Tripartite-GAN contributes to tumor detection. The "upper bound"  
 571 of detection results in the last column/row of these five tables (Table3-7) repre-  
 572 sent the detection results computed directly from ground truth CEMRI images.  
 573 It demonstrates that our synthetic CEMRI images performed close to the real  
 574 CEMRI in tumor detection. The evaluation criterion of accuracy, sensitivity,  
 575 and specificity are defined as follow:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (13)$$

$$\text{Sensitivity : } TPR = \frac{TP}{TP + FN} \quad (14)$$

$$\text{Specificity : } TNR = \frac{TN}{FP + TN} \quad (15)$$

576 where the sensitivity and specificity are equivalent to the true positive rate  
 577 (TPR) and the true negative rate (TNR), respectively. The TP, FP, TN, and  
 578 FN denotes the true positive, false positive, true negative, and false negative  
 579 measurements, respectively.

580 *4.2.8. Adding attention maps into detector improves the performance of tumor  
 581 detection*

582 In order to verify the attention maps have the potential to detect tumors  
 583 and the contribution of adding attention maps into the detector, we perform  
 584 the comparison among three different detectors as follows: 1) Using the atten-  
 585 tion maps instead of feature maps from VGG-16 for the box generation and  
 586 tumor classification by R-CNN. 2) Using the VGG-16-based model to obtain  
 587 feature maps without the help of attention maps for box generation and tumor  
 588 classification by R-CNN. 3) Using the VGG-16-based model to obtain feature  
 589 maps with the help of attention maps for box generation and tumor classifica-  
 590 tion by R-CNN (our detector as shown in Fig.7). The results of the comparison  
 591 are shown in Fig.14, the columns 2, 3, and 4 correspond to the detector 1),  
 592 2) and 3) mentioned above, respectively. It is clear that attention maps have

Table 5: When we defined healthy subjects as positive and hemangioma as negative, the comparison of Tripartite-GAN and three other methods of detection demonstrates that our Tripartite-GAN outperforms three other detection methods. The ablation studies demonstrate that every part of the newly designed Tripartite-GAN contributes to tumor detection.

Method	health(P)/Hemangioma(N)		
	Accuracy	Sensitivity	Specificity
<i>Comparison</i>			
FZM-SVM	78.0±1.3%	87.0±2.6%	92.2±1.4%
Modified Faster-RCNN	78.7±1.0%	87.2±2.2%	92.2±1.4%
U-Net based FCN	79.5±1.1%	89.1±2.1%	92.3±1.2%
<i>Ablation study</i>			
No G + No D	80.3±1.2%	89.4±2.0%	92.3±1.3%
No D	81.1±1.0%	89.4±2.0%	92.4±1.2%
No DAM	83.5±0.8%	91.5±1.8%	94.0±1.0%
No MAM	84.3±0.8%	91.5±1.8%	95.5±1.0%
No GAM	85.8±0.8%	93.6±1.7%	95.6±0.9%
No Atten	87.4±0.7%	95.7±1.5%	95.7±0.8%
No Di-con	89.0±0.6%	95.8±1.6%	97.1±0.8%
No Res-L	90.0±0.7%	95.8±1.5%	98.6±0.9%
<b>Tripartite-GAN</b>	<b>91.3±0.6%</b>	<b>95.9±1.3%</b>	<b>98.6±0.7%</b>
<b>Tripartite-GAN (upper bound)</b>	<b>92.9±0.6%</b>	<b>96.0±1.2%</b>	<b>98.6±0.7%</b>

Table 6: When we defined healthy subjects as positive and HCC as negative, the comparison of Tripartite-GAN and three other methods of detection demonstrates that our proposed Tripartite-GAN outperforms three other detection methods. The ablation studies demonstrate that every part of the newly designed Tripartite-GAN contributes to tumor detection.

Method	health(P)/HCC(N)		
	Accuracy	Sensitivity	Specificity
<i>Comparison</i>			
FZM-SVM	77.9±1.2%	87.0±2.2%	93.9±1.2%
Modified Faster-RCNN	79.5±1.1%	89.1±2.0%	94.8±1.0%
U-Net based FCN	78.9±1.1%	87.2±2.1%	94.0±1.1%
<i>Ablation study</i>			
No G + No D	80.0±1.0%	89.4±1.9%	94.8±1.0%
No D	81.1±0.9%	89.4±1.8%	94.9±0.9%
No DAM	82.6±0.8%	90.0±1.6%	95.0±0.8%
No MAM	83.2±0.8%	89.6±1.7%	95.8±0.7%
No GAM	84.2±0.8%	89.8±1.7%	95.9±0.7%
No Atten	85.3±0.7%	90.0±1.5%	95.9±0.6%
No Di-con	86.8±0.7%	92.0±1.5%	96.7±0.7%
No Res-L	87.4±0.8%	92.0±1.6%	96.8±0.7%
<b>Tripartite-GAN</b>	<b>88.4±0.6%</b>	<b>94.0±1.2%</b>	<b>96.8±0.6%</b>
<b>Tripartite-GAN (upper bound)</b>	<b>90.0±0.6%</b>	<b>96.0±1.1%</b>	<b>97.6±0.5%</b>

Table 7: When we defined hemangioma as positive and HCC as negative, the comparison of Tripartite-GAN and three other methods of detection demonstrates that our proposed Tripartite-GAN outperforms three other detection methods. The ablation studies demonstrate that every part of the newly designed Tripartite-GAN contributes to tumor detection.

Method	Hemangioma(P)/HCC(N)		
	Accuracy	Sensitivity	Specificity
<i>Comparison</i>			
FZM-SVM	78.4±1.2%	84.3±1.7%	82.4±1.5%
Modified Faster-RCNN	79.3±1.0%	84.3±1.5%	83.3±1.5%
U-Net based FCN	79.3±1.1%	85.7±1.5%	83.2±1.5%
<i>Ablation study</i>			
No G + No D	79.8±0.8%	85.7±1.3%	83.3±1.3%
No D	81.2±0.7%	87.1±1.2%	84.8±1.3%
No DAM	83.6±0.7%	88.7±1.2%	86.4±1.2%
No MAM	84.0±0.7%	88.9±1.2%	86.5±1.2%
No GAM	85.0±0.6%	90.3±1.1%	87.2±1.1%
No Atten	85.9±0.7%	91.7±1.1%	88.0±1.0%
No Di-con	87.3±0.6%	91.8±1.0%	88.8±1.1%
No Res-L	88.3±0.6%	91.9±1.0%	89.6±1.0%
<b>Tripartite-GAN</b>	<b>89.2±0.5%</b>	<b>93.2±0.8%</b>	<b>90.3±0.9%</b>
<b>Tripartite-GAN (upper bound)</b>	<b>90.6±0.5%</b>	<b>94.6±0.6%</b>	<b>91.1±0.8%</b>

593 the potential to detect tumors but not accurate. Because attention maps not  
 594 only focus on the tumor but pay more attention to extract all features of all  
 595 anatomy structure in liver MRI for entire liver MRI synthesis. Both the feature  
 596 maps of VGG-16 without attention maps and the feature maps of VGG-16 with  
 597 attention maps focus on tumor information, but the characterizations of tumors  
 598 in VGG-16 with attention maps are more accurate and detailed than without  
 599 attention maps. The detection results of three different detectors demonstrate  
 600 that the attention map added into VGG-16 in the manner of residual connection  
 601 improves R-CNN-based detector to detect the tumor.

## 602 5. Conclusions and Discussion

603 Our proposed Tripartite-GAN successfully synthesized liver CEMRI without  
 604 CA injection on a dataset of 265 subjects. The synthesized liver CEMRI has the  
 605 equivalent value of real CEMRI in clinical diagnosis and then used to accurate  
 606 tumor detection. All of the results demonstrate that Tripartite-GAN can aid in  
 607 the clinical diagnosis as a safe, time-saving, and inexpensive tool to synthesize  
 608 and detect CEMRI without CA injection. The effective application of the hybrid  
 609 convolution, residual learning, and DAM is able to expand the receptive field,  
 610 facilitate convergence, and enhance feature representation. Moreover, it is the  
 611 first that the Tripartite-GAN successfully combined the regular GAN and the

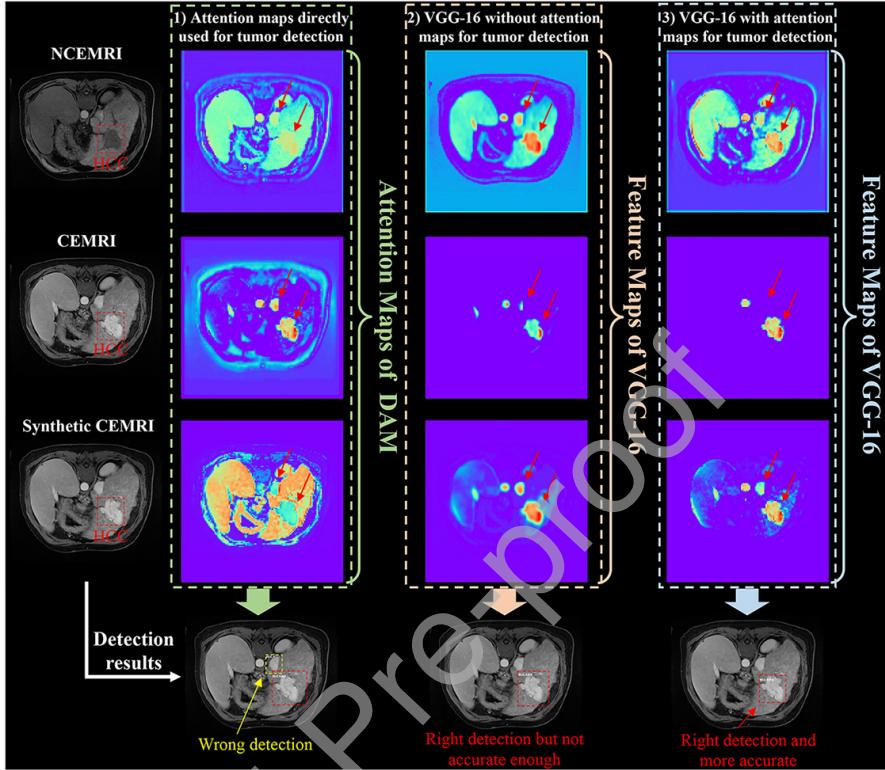


Figure 14: Attention maps not only focus on the tumor but pay more attention to extract all features of all anatomy structure in liver MRI for multi-class liver MRI synthesis. The feature maps of VGG-16 without attention maps are more focused on tumor information. The feature maps of VGG-16 with attention maps also focus on tumor information but more accurate and detailed than without attention maps.

detector. The tripartite loss not only has a stable performance but also achieves that the liver CEMRI synthesis and tumor detection promote each other in an end-to-end framework. Meanwhile, the generator improves the performance of tumor detection via newly adding attention maps into the detector in the manner of residual connection. Besides, the Tripartite-GAN has the potential to synthesize liver CEMRI for tumor segmentation.

#### 618 Acknowledgments

619 This work was funded by the China Scholarship Council (No. 201808370212),  
 620 the National Natural Science Foundation of China (61971271), the Taishan  
 621 Scholars Project of Shandong Province (Tsqn20161023) and the Primary Re-  
 622 search and Development Plan of Shandong Province (No. 2018GGX101018, No.  
 623 2019QYTPY020).

624 **References**

- 625 Akselrod-Ballin, A., Karlinsky, L., Alpert, S., Hasoul, S., Ben-Ari, R., Barkan,  
 626 E., 2016. A region based convolutional network for tumor detection and  
 627 classification in breast mammography, in: Deep learning and data labeling  
 628 for medical applications. Springer, pp. 197–205.
- 629 Beers, A., Brown, J., Chang, K., Campbell, J.P., Ostmo, S., Chiang, M.F.,  
 630 Kalpathy-Cramer, J., 2018. High-resolution medical image synthesis us-  
 631 ing progressively grown generative adversarial networks. arXiv preprint  
 632 arXiv:1805.03144 .
- 633 Ben-Cohen, A., Klang, E., Raskin, S.P., Soffer, S., Ben-Haim, S., Konen, E.,  
 634 Amitai, M.M., Greenspan, H., 2019. Cross-modality synthesis from ct to  
 635 pet using fcn and gan networks for improved automated lesion detection.  
 636 Engineering Applications of Artificial Intelligence 78, 186–194.
- 637 Chongxuan, L., Xu, T., Zhu, J., Zhang, B., 2017. Triple generative adversarial  
 638 nets, in: Advances in neural information processing systems, pp. 4088–4098.
- 639 Costa, P., Galdran, A., Meyer, M.I., Niemeijer, M., Abràmoff, M., Mendonça,  
 640 A.M., Campilho, A., 2017. End-to-end adversarial retinal image synthesis.  
 641 IEEE transactions on medical imaging 37, 781–791.
- 642 Digumarthy, S.R., Sahani, D.V., Saini, S., 2005. Mri in detection of hepatocel-  
 643 lular carcinoma (hcc). Cancer Imaging 5, 20.
- 644 Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y., 2017. Automatic brain tumor  
 645 detection and segmentation using u-net based fully convolutional networks,  
 646 in: annual conference on medical image understanding and analysis, pp. 506–  
 647 517.
- 648 Emami, H., Dong, M., Nejad-Davarani, S.P., Glide-Hurst, C.K., 2018. Generat-  
 649 ing synthetic cts from magnetic resonance images using generative adversarial  
 650 networks. Medical physics 45, 3627–3636.
- 651 Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., Greenspan,  
 652 H., 2018. Gan-based synthetic medical image augmentation for increased cnn  
 653 performance in liver lesion classification. Neurocomputing 321, 321–331.
- 654 Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H., 2019. Dual attention  
 655 network for scene segmentation, in: Proceedings of the IEEE Conference on  
 656 Computer Vision and Pattern Recognition, pp. 3146–3154.
- 657 Girshick, R., 2015. Fast r-cnn, in: Proceedings of the IEEE international con-  
 658 ference on computer vision, pp. 1440–1448.
- 659 Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair,  
 660 S., Courville, A., Bengio, Y., 2014. Generative adversarial nets, in: Advances  
 661 in neural information processing systems, pp. 2672–2680.

- 662 Halavaara, J., Breuer, J., Ayuso, C., Balzer, T., Bellin, M.F., Blomqvist,  
 663 L., Carter, R., Grazioli, L., Hammerstingl, R., Huppertz, A., et al., 2006.  
 664 Liver tumor characterization: comparison between liver-specific gadoxetic  
 665 acid disodium-enhanced mri and biphasic ct-a multicenter trial. Journal of  
 666 computer assisted tomography 30, 345–354.
- 667 He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image  
 668 recognition, in: Proceedings of the IEEE conference on computer vision and  
 669 pattern recognition, pp. 770–778.
- 670 Hu, H., Gu, J., Zhang, Z., Dai, J., Wei, Y., 2018a. Relation networks for object  
 671 detection, in: Proceedings of the IEEE Conference on Computer Vision and  
 672 Pattern Recognition, pp. 3588–3597.
- 673 Hu, J., Shen, L., Sun, G., 2018b. Squeeze-and-excitation networks, in: Proceed-  
 674 ings of the IEEE conference on computer vision and pattern recognition, pp.  
 675 7132–7141.
- 676 Idée, J.M., Port, M., Raynal, I., Schaefer, M., Le Greneur, S., Corot, C., 2006.  
 677 Clinical and biological consequences of transmetallation induced by contrast  
 678 agents for magnetic resonance imaging: a review. Fundamental & clinical  
 679 pharmacology 20, 563–576.
- 680 Korkinof, D., Rijken, T., O'Neill, M., Yearsley, J., Harvey, H., Glockner, B.,  
 681 2018. High-resolution mammogram synthesis using progressive generative  
 682 adversarial networks. arXiv preprint arXiv:1807.03401 .
- 683 Low, R.N., 2007. Abdominal mri advances in the detection of liver tumours and  
 684 characterisation. The Lancet Oncology 8, 525–535.
- 685 Mahasseni, B., Lam, M., Todorovic, S., 2017. Unsupervised video summariza-  
 686 tion with adversarial lstm networks, in: Proceedings of the IEEE conference  
 687 on Computer Vision and Pattern Recognition, pp. 202–211.
- 688 Marckmann, P., Skov, L., Rossen, K., Dupont, A., Damholt, M.B., Heaf, J.G.,  
 689 Thomsen, H.S., 2006. Nephrogenic systemic fibrosis: suspected causative  
 690 role of gadodiamide used for contrast-enhanced magnetic resonance imaging.  
 691 Journal of the American Society of Nephrology 17, 2359–2362.
- 692 Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. arXiv  
 693 preprint arXiv:1411.1784 .
- 694 Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q.,  
 695 Shen, D., 2018. Medical image synthesis with deep convolutional adversarial  
 696 networks. IEEE Transactions on Biomedical Engineering 65, 2720–2730.
- 697 Peng, C., Zhang, X., Yu, G., Luo, G., Sun, J., 2017. Large kernel matters—  
 698 improve semantic segmentation by global convolutional network, in: Pro-  
 699 ceedings of the IEEE conference on computer vision and pattern recognition,  
 700 pp. 4353–4361.

- 701 Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time  
 702 object detection with region proposal networks, in: Advances in neural information  
 703 processing systems, pp. 91–99.
- 704 Shen, Y., Luo, P., Yan, J., Wang, X., Tang, X., 2018. Faceid-gan: Learning a  
 705 symmetry three-player gan for identity-preserving face synthesis, in: Proceedings  
 706 of the IEEE Conference on Computer Vision and Pattern Recognition,  
 707 pp. 821–830.
- 708 Simon, V., Bert, D.B., Davy, N., Luc, V.G., 2019. A three-player GAN:  
 709 generating hard samples to improve classification networks. arXiv preprint  
 710 arXiv:1903.03496 .
- 711 Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-  
 712 scale image recognition. arXiv preprint arXiv:1409.1556 .
- 713 Singh, A., et al., 2015. Detection of brain tumor in mri images, using combi-  
 714 nation of fuzzy c-means and svm, in: 2015 2nd International Conference on  
 715 Signal Processing and Integrated Networks (SPIN), IEEE. pp. 98–102.
- 716 Son, J., Park, S.J., Jung, K.H., 2017. Retinal vessel segmentation in fundoscopic  
 717 images with generative adversarial networks. arXiv preprint arXiv:1706.09318  
 718 .
- 719 Springenberg, J.T., 2015. Unsupervised and semi-supervised learning with cat-  
 720 egorical generative adversarial networks. arXiv preprint arXiv:1511.06390 .
- 721 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N.,  
 722 Kaiser, L., Polosukhin, I., 2017. Attention is all you need, in: Advances in  
 723 neural information processing systems, pp. 5998–6008.
- 724 Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic  
 725 demons: Efficient non-parametric image registration. NeuroImage 45, S61–  
 726 S72.
- 727 Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., Cottrell, G., 2018.  
 728 Understanding convolution for semantic segmentation, in: 2018 IEEE winter  
 729 conference on applications of computer vision (WACV), IEEE. pp. 1451–1460.
- 730 Wolterink, J.M., Dinkla, A.M., Savenije, M.H., Seevinck, P.R., van den Berg,  
 731 C.A., Işgum, I., 2017. Deep mr to ct synthesis using unpaired data, in:  
 732 International Workshop on Simulation and Synthesis in Medical Imaging,  
 733 Springer. pp. 14–23.
- 734 Yang, W., Zhong, L., Chen, Y., Lin, L., Lu, Z., Liu, S., Wu, Y., Feng, Q., Chen,  
 735 W., 2018. Predicting ct image from mri data through feature matching with  
 736 learned nonlinear local descriptors. IEEE transactions on medical imaging  
 737 37, 977–987.

- 738 Zhang, H., Goodfellow, I., Metaxas, D., Odena, A., 2018. Self-attention gener-  
739 ative adversarial networks. arXiv preprint arXiv:1805.08318 .
- 740 Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid scene parsing  
741 network, in: Proceedings of the IEEE conference on computer vision and  
742 pattern recognition, pp. 2881–2890.
- 743 Zhao, S., Wu, X., Chen, B., Li, S., 2019. Automatic spondylolisthesis grad-  
744 ing from mrис across modalities using faster adversarial recognition network.  
745 Medical Image Analysis , 101533.

**Declaration of Competing Interest**

none.