原理：心脏跳动会 使皮肤中的血容量呈周期性变化，其中的血红蛋白含量呈周期性变化，对红光的吸收也呈周期性变化，反射给摄像头的红光强度便呈周期性变化，所以呈现出的最终变化为**皮肤颜色随心跳的周期性变化**。

目的：从皮肤微弱的周期性颜色变化中提取心动周期。

Dual-GAN: Joint BVP and Noise Modeling for Remote Physiological Measurement

**研究现状：**现有的方法大都关注**如何增强或提取人脸视频中微弱的BVP信号，**而很少**对**人脸视频内容占主要成分的**噪声进行建**模的。这会导致

解决方法：提出了一种 对偶学习的生成对抗网络（Dual GAN），由两部分组成：

STMap ----深度网络---> BVP signal

ROI **Alignment？** and Fusion Block

图片包含 图表

描述已自动生成

Alignment

问题1：

（从对时间序列*L*做卷积的维度）由于人脸不是一个an ideal Lambertian object（Lambertian表面是指在一个固定的照明分布下从所有的视场方向上观测都具有相同亮度的表面，） 所以，不同的ROI的BVP信号和噪声分布也是不同的。

而现有的方式对每个ROI提取的时间序列，都用相同的时域卷积来提取BVP信号。

解决办法：

对每个ROI都分别进行不同的卷积。

图示

描述已自动生成

问题2：

（从对ROIs *N*做卷积的维度）不同ROI中包含的BVP信号应当是几乎同步的。如果ROI的感受野很大能够统筹兼顾所有的ROI那一定能取得很好的效果。

而现有的方式是直接对STMap做卷积，每次做卷积只能汇聚到相邻的几个ROI的信息。

解决方法：

图示

描述已自动生成

图示

描述已自动生成

通过引入通道注意力机制，自行判断哪个ROI的哪个颜色通道（一共是N×C个时间序列）能够提取到有用的信息。提取特征时能同时对所有的ROI的所有通道统筹兼顾。

**整个block下来输入输出的尺寸是不变的。**

Fb模块： 从STMap提取BVP信号

图示

描述已自动生成

具体结构如下：

图片包含 图示

描述已自动生成

Auxiliary model分支直接从stmap映射为心率（反正有数据集），旨在用多任务的方式来提升模型的鲁棒性

Fn模块：生成“合成的STMap”

真实BVP信号 ----Gphy----> 生成的无噪声STMap

纯噪声 ----Gnoise---> 生成的STMap的噪声

生成的无噪声STMap + 生成的STMap的噪声 = 合成的STMap

图示

描述已自动生成

表格

描述已自动生成

现在我们有 生成器Fb 从m（真实的STMap） 生成s est（虚假的BVP信号 ）

生成器Fn 从sgt（真实的BVP信号） 生成虚假的m syn（STMap）

两个生成模块 共用 **一个判别器D**

判别器D输入为 STMap和BVP signal的concatenate，具体方式如下：

图片包含 表格

描述已自动生成

那么我们可以用 真实的stmap，虚假的stmap，真实的bvp信号，虚假的bvp信号共构成三种搭配。

真实的stmap & 真实的bvp信号 送入判别器，使其能够正确建立stmap和bvp信号的关系

对于BVP Modeling GAN：真实的stmap & 虚假的bvp信号 送入判别器，旨在训练其能够正确将 stmap映射为 bvp信号的能力

对于Noise modeling GAN：虚假的stmap & 真实的bvp信号 送入判别器，旨在训练其能够正确对噪声进行建模，能够从真实bvp信号合成其对应的STMap的能力。

判别器判断并非判断输入真假，而是判断所输入的stmap和bvp信号两者是否匹配，标签就是以是否匹配来给的。

文本, 信件

描述已自动生成

噪声的建模能力：

至此还没完，因为我们要的时从stmap到bvp的fb，所以另一对噪声进行建模的fn对fb的制约仅能通过判别器而且感觉基本没有起到这种指导的作用。感觉噪声建模的那部分没必要。

然后作者又加入了最后一条分支，将fn和fb连接起来，用fn学会的从bvp映射至stmap的能力，指导fb对真实的stmap去噪但不去有效信号（合理地滤波）地映射至bvp，即训练bvp的滤波能力。

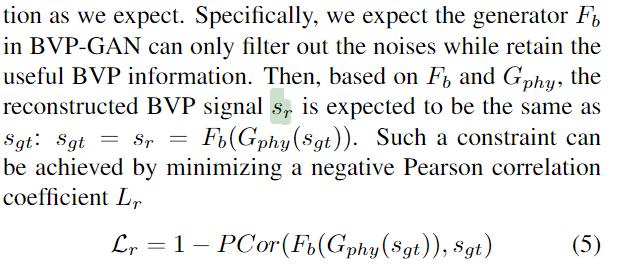
（即将gphy从真实的bvp映射成的stmap，应当经fb得到的bvp和原bvp相同，不同就产生梯度）

文本

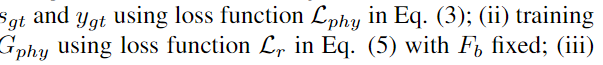
描述已自动生成

我服了，写的这个小分支训练Gphy，Fb固定

明明说



然后，



---------------------------------------------------------------------------

觉得这个网络比较耐啃的一点：

我觉得有一点问题是，因为全文在强调对噪声的建模能力，但是最后也没有用Gnoise做什么，比如用真实的stmap减去这个生成的noise得到的纯净stmap，和gphy从bvp映射成的不含噪声的stmap，这两个stmap 考察他俩的皮尔逊相关系数，

当然或许这样计算量更大，而且只能说是对Gnoise和Gphy优化而没有优化我们所需的那个Fb。所以作者采用了刚才那种方式将两个gan建立关系。

而用那种方法建立关系又不能直观地体现出来，对噪声建模的作用，而顶多说是，Fb的去噪能力很好，Gphy将bvp信号转化为mphy的能力很好。

硬要说Gnoise的作用，起到指导作用也可以扯上关系，比如

如果Gnoise效果很好，其能将噪声的高斯分布映射成stmap的噪声分布，能够很好地对噪声建模，然后和Gphy生成的stmap相加获得合成的stmap，如果此时这个stmap很差判别器将其判断为不匹配，那肯定是Gphy的问题它不能够将bvp很好地转化为stmap，便会对Gphy进行优化，使其能更好地将bvp很好地转化为stmap，便通过这个分支指导Fb的滤波能力。

或者说

Fn分支训练的很好的话，gnoise能够产生完全的噪声（能够很好地对噪声建模），且gphy能够精准地将bvp转化为stmap，而它们之所以能分的这么开，便得益于判别器提供的梯度。

Gnoise建模的好，使得gphy映射的很准确，ghpy继而指导fb滤波。

应当说fb的好也得益于很好的噪声建模能力。

但是扯得很远。

不过，如果Gnoise的能力真的很好，我们还可以将它冻结的方式，应该能将Gphy训练到极致。

不过不管怎么说，只要连接起来了，两个GAN的耦合能力就大大提升了，能建立起很复杂的关系，所以只要有足够的算力，效果肯定是非常好的。