

# 电磁信号预处理

计算 1811 罗廷杨

## 1、概述

Conv-Tasnet 是一种基于监督学习的深度神经网络方法，它实现盲源分离的实验流程主要有两个阶段：训练阶段和测试阶段，训练阶段将准备好的训练集放入深度神经网络进行训练得出模型，并用验证集验证模型的准确性，测试阶段使用训练阶段获得的模型来对测试集进行分离。基于监督学习的流程如图所示：

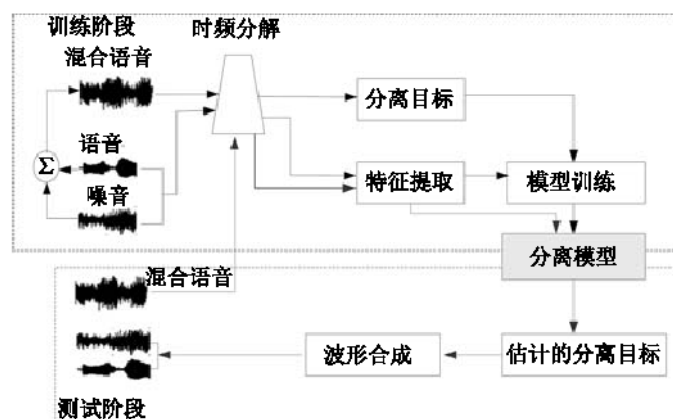


图 1 基于监督学习的流程图

可见我们要使用监督学习的方法来进行盲源分离实验第一步就是数据集的制作，由于训练集、验证集、测试集的制作原理一样，下面就以训练集的制作为例。

首先我来看看语音训练集的制作，语音训练集里面包含三种信号：语音 1（s1）、语音 2（s2）、混合语音（mix）。由于监督学习中需要将训练集拿给深度神经网络进行训练，所以应该尽量的满足：mix 只由 s1 和 s2 混合而成，不包含其他的干扰元素，即  $mix = s1 + s2$ 。由这样的三种信号就可以构成语音的训练集。

了解了语音信号训练集的制作方式之后，我们将它的训练集结构推广到电磁信号训练集上，电磁信号的训练集也有三种信号：信号 1（s1）、信号 2（s2）、混合信号（mix），并且应该尽量的满足 mix 只由 s1 和 s2 混合而成。我们秉持这个原则去采集数据并对数据进行处理来构造我们的电磁信号训练集。

## 2、数据集采集

### 2.1 设备介绍

实验室的设备有两个采集口，一次可以采集两路数据，这两路数据的来源是：辐射信号和 VGA 信号，

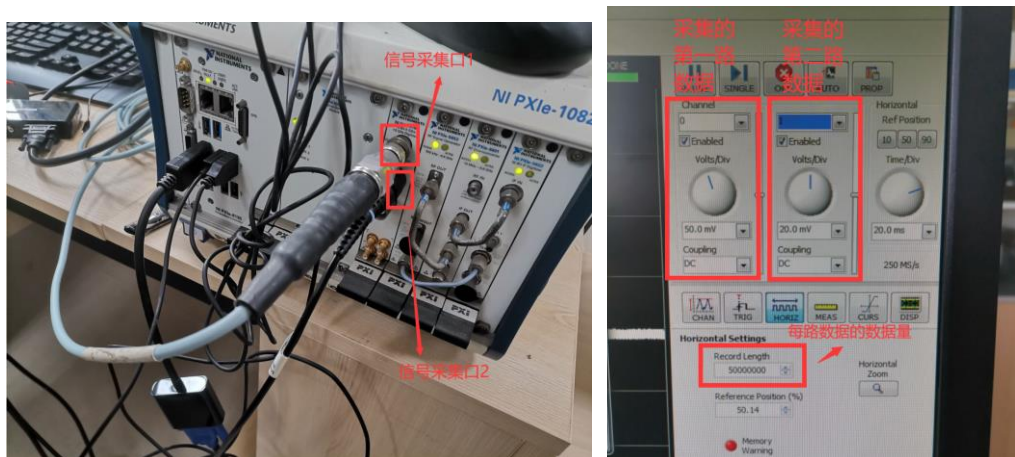


图 2 采集设备的结构图



图 3 采集电磁辐射信号设备



图 4 采集 VGA 信号设备

VGA 各个接口所采集到的信号含义如下：

## 接口 管脚定义



DB-15VGA连接器	针脚	说明	针脚	说明	针脚	说明
	1	红色	6	红色回路	11	空（地址码）
	2	绿色	7	绿色回路	12	SDA (DDC数据)
	3	蓝色	8	蓝色回路	13	水平同步（行）
	4	空（地址码）	9	空（地址码）	14	垂直同步（场）
	5	接地（模拟）	10	接地（同步回路）	15	SCL时钟

图 5 VGA 信号含义图

## 2.2 采集信号介绍

了解了实验的设备之后，我们对数据进行采集，采集的数据如下：

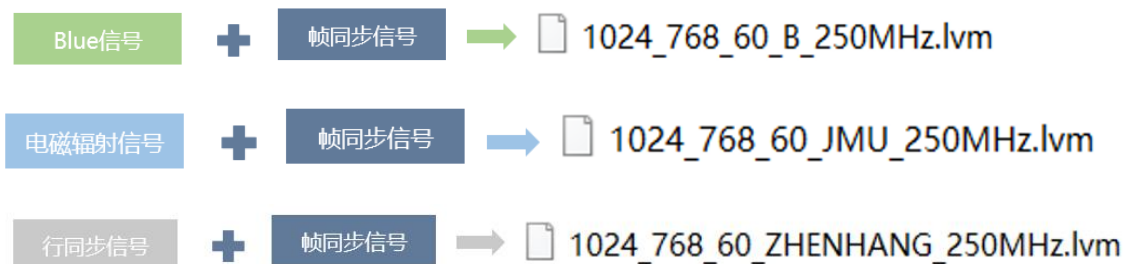


图 6 采集信号结构图

Blue 信号：VGA 信号中的蓝信号

电磁辐射信号：用卡钳采集到的信号

帧同步信号：VGA 信号中的垂直同步信号，用来标识显示屏中每帧信号的开始与结束

行同步信号：VGA 信号中的水平同步信号，用来标识显示屏中每行的开始与结束

关于帧同步信号与行同步信号的进一步理解，可以看图 7：

帧同步信号就是图中的 VSync，行同步信号就是图中的 HSync，。假设显示屏中的内容是“集美大学”字样，那么一帧的信号就是一屏的“集美大学”字样，我们之所以能看到屏幕一直显示“集美大学”是因为显示屏一直在刷这个“集

美大学”帧，这样当我们采集数据的时候可能一下就采集了很多帧“集美大学”，我们无法确定每一帧是从哪里开始和结束，所以这里就引入了帧同步信号来确定我们采集到的数据每一帧开始和结束位置。而对于一个屏幕来说它是一个矩形的形状，它有很多行，因此，“集美大学”帧信号要进行准确的显示，就必须知道帧信号中每一行的开始位置和结束位置，这就需要通过行同步信号进行标定。由此可见帧同步信号和行同步信号就是起一个定位作用。

细心的同学会发现，每次的采集里面都会含有一路帧同步信号，而行同步信号只采集了一次，这是因为每种信号都要确定帧的开始与结束，所以，当我们在采集信号时，就一定要和帧同步信号一起采集，而由于我们的显示屏是固定的，所以行数和行长度就是不变的，所以只需采集一次行同步信号即可。

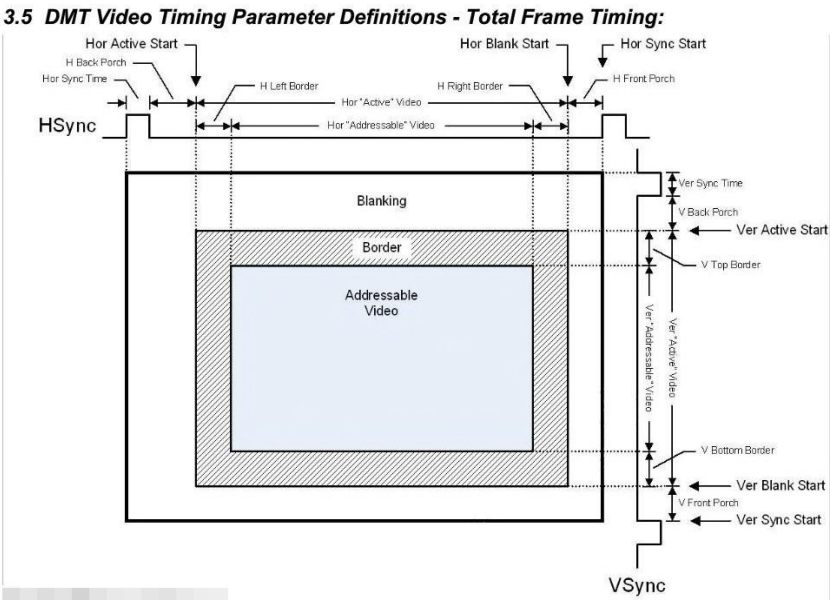


图 7 显示屏结构图

### 2.3 采集信号的复现

在采集完数据后，可以使用 uedit64.exe 软件打开数据看看里面的结构，下面以打开 1024\_768\_60\_JMU\_250MHz.lvm 这个文件为例，左边一列是 Chan 0 其数据为信号采集口 1 采集的数据，右边一列是 Chan 1 其数据为信号采集口 2 采集的数据：

31	***End_of_Header***			
32	X_Value	Chan 0	Chan 1	Comment
33	-0.00091	4.68897		
34	-0.00023	4.68897		
35	0.00046	4.71195		
36	-0.00137	4.75792		
37	0.00046	4.71195		
38	0.00023	4.71195		
39	-0.00046	4.71195		
40	0.00114	4.75792		

电磁辐射信号
帧同步信号

图 8 1024\_768\_60\_JMU\_250MHz.lvm 文件结构图

为了能够更加直观的看到采集到的数据，取出每一路的数据，然后使用 Matlab 程序 Re\_Image.m 进行复现，由于帧同步信号和行同步信号只是起到定位作用这里不对它们进行复现，只对 Blue 信号和电磁辐射信号进行复现。

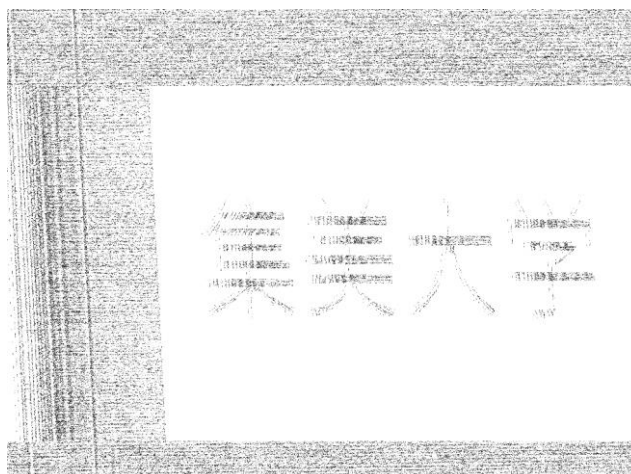


图 10 Blue 信号复现图

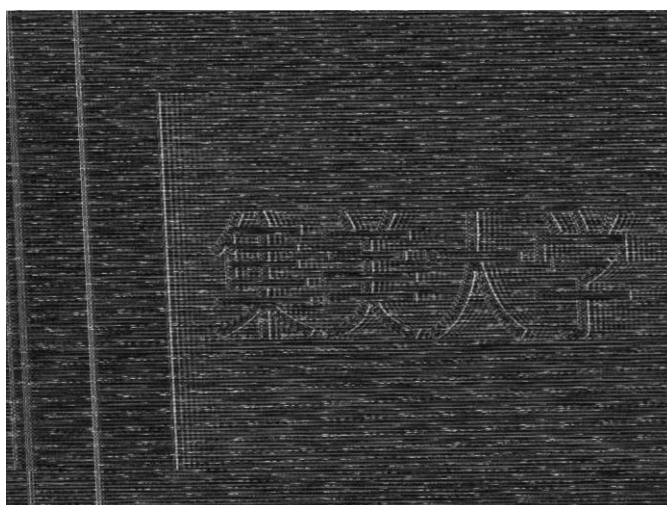


图 11 电磁辐射信号复现图



### 3、电磁信号数据集制作

看到这两个信号后，我们应该选取什么来充当训练集中  $s_1$ 、 $s_2$  和  $mix$  呢？我们现在可以知道的是最终训练出来的模型是为了分离电磁信号，也就是类似上面的电磁辐射信号，因此我们可以将电磁辐射信号作为  $mix$ ，那  $s_1$ 、 $s_2$  该如何获得呢？根据前面的分析，我们知道为了使模型的训练效果更好应该尽量的满足： $mix$  只由  $s_1$  和  $s_2$  混合而成，不包含其他的干扰元素，即  $mix = s_1 + s_2$ 。这样我们就有了一个想法，能不能直接从  $mix$  信号中直接分出  $s_1$  和  $s_2$  呢。这就让我们想到使用掩码，我们只需要构造出  $s_1$  的掩码  $mask_1$  和  $s_2$  的掩码  $mask_2$ ，将  $mask_1$  点乘  $mix$  就可以得到  $s_1$ ，将  $mask_2$  点乘  $mix$  就可以得到  $s_2$ 。

这样一来我们只需将注意力放在如何构造掩码身上，下面对掩码的构造方法进行介绍。已知我们能够很容易的采集到“集美大学”帧中的蓝信号（Blue 信号），而电磁辐射信号中也会包含蓝信号，那么我们就可以尝试着获取 Blue 信号有效数据的位置来制作电磁辐射信号中蓝信号的掩码，再从电磁辐射信号中提取出蓝信号作为  $s_1$ ，这样我们就获取信号  $s_1$ 。下面我们先来看看如何获取 Blue 信号有效数据的位置。

为了方便大家理解，这里给出一个更为简单的位置获取的例子。如何获取下图中有效数据的起始位置和结束位置呢？

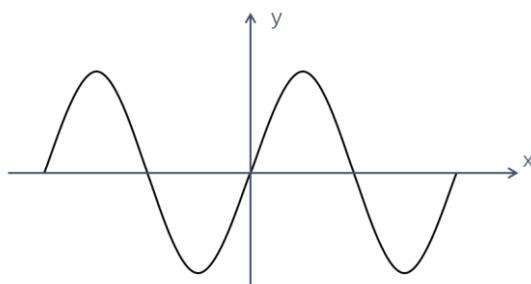


图 12 实例信号图

为了解决这个问题，首先要知道什么是有效数据，所谓的有效数据需要你选定一个标准，当符合这个标准时就是有效数据，不符合这个标准时就是无效数据。知道这个后，我们就可以通过二值化和差分来解决这个问题。

假设这里设定的标准是  $y$  大于 0 的数据就是有效数据，那么二值化操作中，取的标准就是  $y=0$ ，然后进行划分操作， $y>0$  时令  $y=1$ ， $y\leq 0$  时令  $y=0$ ：

1.取标准： $y=0$

2.划分： $\begin{cases} y>0 \text{ 时, } y=1 \\ y\leq 0 \text{ 时, } y=0 \end{cases}$

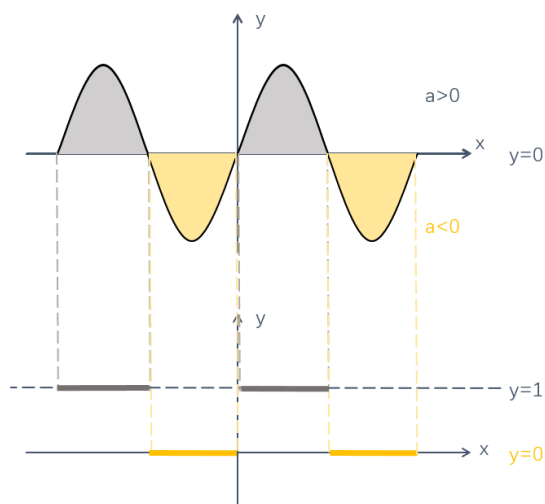


图 13 二值化操作流程

二值化之后再对其进行差分操作，即将“前面的值减去后面的值”或者将“后面的值减去前面的值”，这里使用前者进行操作，这样一来值为-1 的地方为有效数据的起始位置，值为 1 的地方为有效数据的结束位置

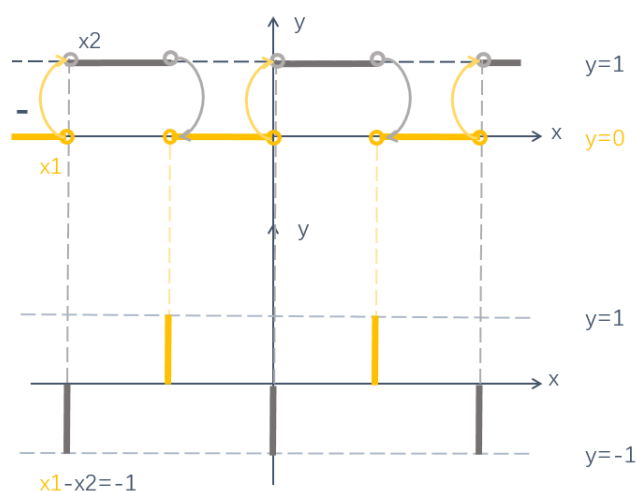


图 14 差分操作流程

我们可以使用上述方法来获取 Blue 信号有效数据的起始和结束的位置，其中值得注意的是差分标准的选取，需要根据信号的结构来设置，这里选取的标准

是信号的极差。

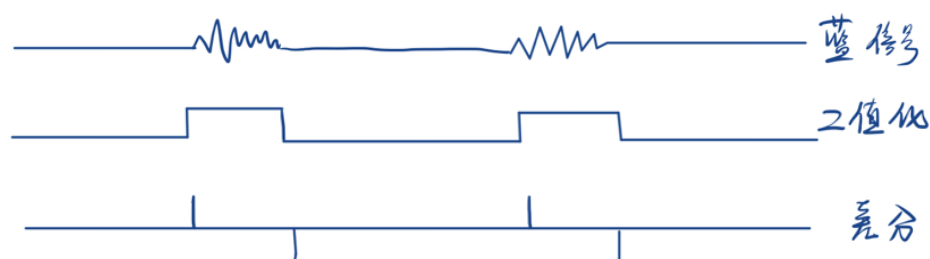


图 15 位置获取流程图

获取 Blue 信号有效数据的起始和结束位置之后，就可以利用这个来制作 Blue 信号的掩码了。制作过程为：先取绝对值，然后再在每个值为 1 的位置替换 10 个点（替换的点数不固定）。为什么需要替换 10 个点呢？因为要用这 10 个点做一个有效数据的扩展，从而使得经过掩码处理后得到的 s1 信号中有更多的有效数据，提高复现的可辨识度。

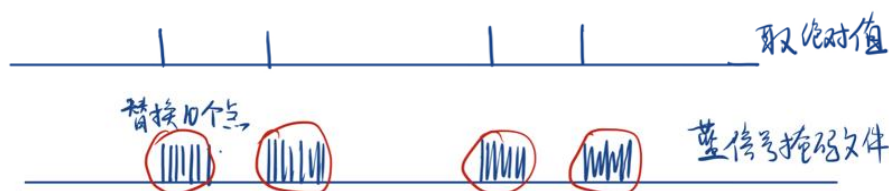


图 16 掩码制作流程图

有了这个通过 Blue 信号制作出来的掩码文件，将其与 mix 信号做点乘就可以获得训练集中的 s1 信号。

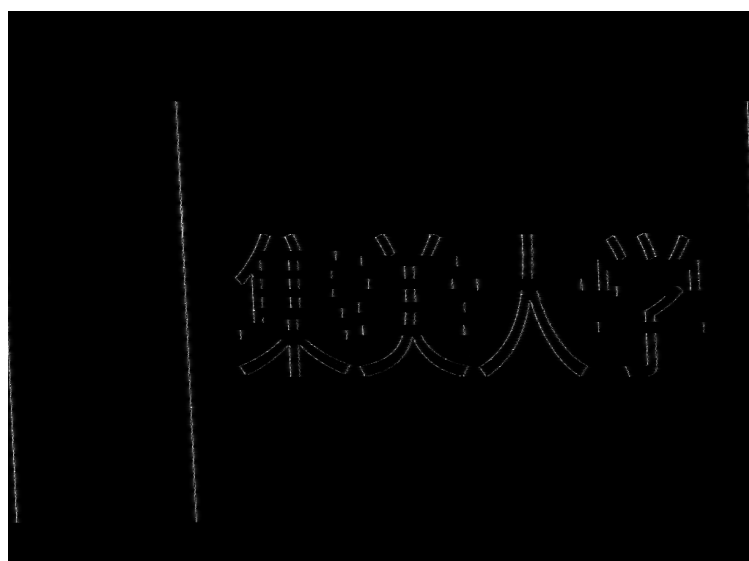


图 17 s1 信号复现图

获取 s1 后，如何获取 s2 呢？我们可能会直接想到用 mix 减去 s1 就是 s2



了，来看看这样处理后的结果。

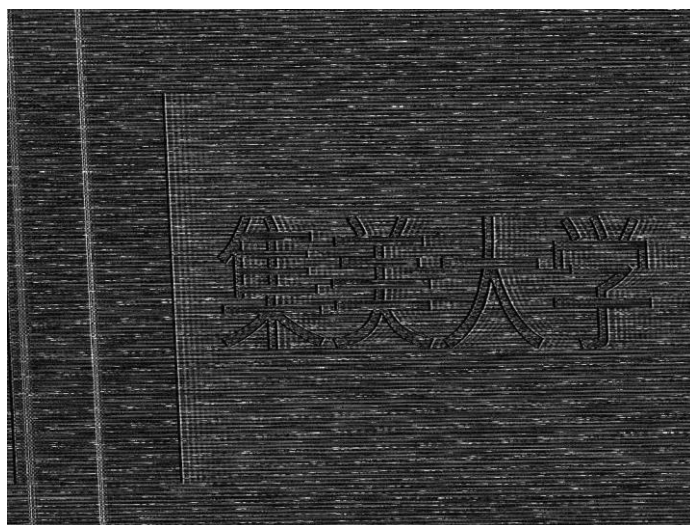


图 18 mix 减去 s1 信号复现图

可以看到结果中存在着去除 s1 后的空白和剩余的集美大学字样，为了保证训练的效果这里使用周围的噪声点来替换空白和集美大学字样。这样就构造出了近似的 s2，当然这样必然存在误差！

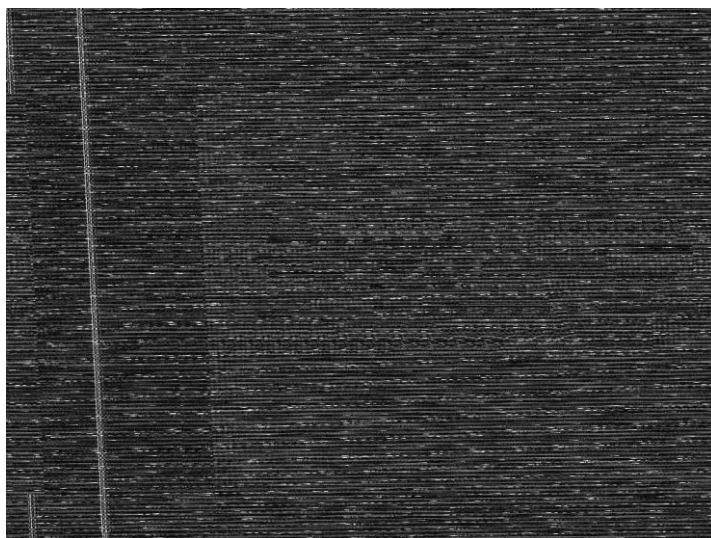


图 19 s2 信号复现图

到这里电磁信号的训练集就制作完成了，我们按照这个思路可以分别再制作出验证集和测试集（注意训练集、验证集、测试集中的数据不能相同），这样电磁信号的预处理结束。