

به نام خدا



درس: علوم اعصاب محاسباتی مقدماتی

استاد: دکتر حمید کربلایی آقاجان

گزارش پروژه ۴

سید محمد امین منصوری طهرانی

۹۴۱۰۵۱۷۴

علی شیرالی

۹۴۱۰۹۱۶۵

قسمت اول: آشنایی با مقاله پژوهش اصلی

۱. هدف مقاله بررسی رابطه رفتار موسیقایی مغز با ساختار هندسی واکسل‌های فعال هنگام گوش دادن به موسیقی است. همان‌طور که در مقاله اشاره شده پیشرفت‌های جزئی در بدست آوردن پاسخ‌های tuning در قسمت‌های اولیه شنیداری بدست آمده است ولی در مورد پاسخ به تحریک‌های دسته‌بندی شده (categorical) مثل تحریک با ژانرهای مشخص و ارتباط آن‌ها با واکسل‌های فعال در بیرون قسمت اولیه شنیداری بررسی‌ای اتفاق نیفتاده‌است. ادعای این پژوهش این است که رمزنگاری نوروئی یا همان چیزی که به عنوان neural encoding با آن آشنا هستیم برای صفات مجزای موسیقی در آرایش چند واکسله نمایان می‌شود. این واکسل‌ها ابعاد فضایی از فیچرها را تشکیل می‌دهند که تعدادی از این فیچرها در کنار هم که همان گروه‌های واکسلی مورد نظر مقاله است قادرند رفتار موسیقایی مغز را توصیف کنند. همچنین ادعای دیگر این پژوهش این است که این decoding طیف وسیعی از ویژگی‌ها و رفتارهای موسیقایی را از دسته‌بندی ژانرها که بسیار کلی است تا تفاوت ملودی‌ها که تفاوت‌های جزئی در فرکانس‌های اصلی (pitch) دارند را توصیف می‌کند. روشی که از آن بهره می‌گیرد نیز Multivoxel Pattern Analysis است و دسته‌ای از واکسل‌ها را در پاسخ به تحریک خاصی در نظر می‌گیرد.

۲. خود مقاله داده‌ها را نگرفته است و از داده‌های آزمایش دیگری استفاده کرده‌است. در این آزمایش ۲۰ شخص راست دست شامل ۱۲ آقا و ۸ خانم به صورت کاملاً داوطلبانه انتخاب شده‌اند. طبق گزارش همه شنوایی معمولی بدون مشکل موقت یا دائمی داشتند و همچنین تجربه اختلال نورولوژیکی نداشتند. اطلاعات شغل و تحصیلات و ترجیحات موسیقایی آن‌ها نیز در گزارش گرفته شده‌است.

آزمایش بدین صورت است که ۲۵ قطعه موسیقی از ۵ ژانر مختلف انتخاب شده است و هر کدام از یکی هنرمندان شاخص آن ژانر انتخاب شده‌است. ژانرها به این شرح هستند: ۱. صدای محیط ۲. کانتری ۳. متال ۴. Rocknroll ۵. سمفونیک. هر کدام از این ۲۵ قطعه ۶ ثانیه طول می‌کشد. اول و آخر صداها نیز با اضافه شدن سینوسی‌های ناقص زیادشونده هموار شده‌است. تفاوت اصلی قطعه‌های موسیقی داشتن/نداشتن صدای سازهای کوبه‌ای یا آواز انسان است. از هر شخصی ۸ بار تست گرفته می‌شود و در هر ران همان قطعه‌های ۶ ثانیه‌ای منتهی با ترتیب متفاوت پخش می‌شوند. فاصله زمانی بین قطعه‌ها نیز عددی از اعداد ۸-۶-۴ ثانیه به صورت تصادفی بود. در هر ژانر نیز یک‌بار یک سوال برای نگه‌داشتن توجه شخص به آزمایش در مورد آن موسیقی پرسیده می‌شد و اطلاعات این موسیقی دور ریخته می‌شود.

۳. داده‌های fmri در حین تحریک صوتی برداشته می‌شدند و اسکن‌های وزن‌دار T1 با ۳ تسلا و اسکن‌های وزن‌دار T2 با ۷ تسلا برای قطعه‌هایی از مغز که قسمتی از آن‌ها پوشش داده می‌شد گرفته شدند. در هر ران ۱۵۳ حجم (ترکیب عکس‌های مغز در کنار هم یک حجم را بازسازی می‌کند). تولید شده‌است. رزولوشن

زمانی و یا همان فاصله بین گرفتن این حجم‌ها که از محدودیت سرعت دستگاه ناشی می‌شود ۲ ثانیه بوده‌است. هر حجم شامل ۳۶ برش از مغز است که هر برش ۱/۴ میلی‌متر ضخامت دارد. مساحت هر واکسل نیز $1/4 * 1/4$ است. بین برش‌ها نیز فضای خالی به میزان ۱۰ درصد در نظر گرفته شده تا تعادلی رزولوشن فضایی و همچنین حجم پوشش داده شده و زمان گرفتن این حجم‌ها برقرار شود. اصلاح حرکت سر در داده‌ها نیز به صورت آنلاین انجام می‌شد. داده‌های گرفته شده با مدل GLM برای هر تحریک به ویژگی‌های هر واکسل نگاشته می‌شدند.

۴. قبل از کلسیفای کردن بخشی از داده‌ها برای cross validation کنار گذاشته می‌شوند. برای کلسیفایرهای آهنگ از کل یک ران به عنوان validation استفاده شده و برای کلسیفایرهای ژانر یک قطعه از هر ژانر در تمام ران‌ها جدا به عنوان validation نگه داشته می‌شود. برای کلسیفای کردن نیز از روش CSVM استفاده شده است. ۳ ناحیه مورد علاقه (Region of interest) از یک اطلس انتخاب شدند و این انتخاب بر اساس دانش قبلی در مورد نقش این ناحیه‌ها در دسته‌بندی موسیقی انجام شد. همچنین برای کاهش اثر واکسل‌های نویزی در عملکرد کلسیفایر الگوریتم انتخاب فیچرها بر اساس یک حساسیت فقط ۵۰۰۰ واکسل موثر اول در هر ROI را نگه داشت. این کار در واقع به وسیله ANOVA انجام شد که در قسمت چهارم جزئیات آن بیشتر مورد بحث قرار می‌گیرد. در واقع یک آزمون فرض است که در مساله ما مشخص می‌کند پاسخ یک واکسل در برابر تحریک‌های مختلف به چه میزان تمایز دهنده است. آنالیز ROI نیز به بررسی خوشه‌هایی از واکسل‌ها در ناحیه‌های فعال مورد نظر می‌گویند.

۵. نتیجه confusion matrix برای این روش کلسیفای کردن (SVM) به این ترتیب است که اتفاق misclassify شدن با احتمال بیشتری برای قطعه‌های هم ژانر و یا «هم‌صدا» رخ می‌دهد و این هم‌صدایی به معنی داشتن سازهای از جنس مشترک و یا داشتن صدای انسان تعبیر می‌شود. همچنین دقت‌های فعالیت رفتاری از کلسیفایر SVM بیشتر است.

۶. نتایج کلسیفای کردن موسیقی‌ها از گزارش‌های قبل خیلی بهتر بوده و علت آن می‌تواند استفاده از دستگاه دقیق‌تر بوده باشد که رزولوشن فضایی را دو برابر می‌کند و تشخیص الگوهای واکسل‌ها را با جزئیات بیشتری فراهم می‌سازد.

کلسیفای کردن ژانرها دقتی کمتر از پژوهش‌های قبلی ذکر شده در مقاله دارد اما cross validation دقیق‌تری دارد. این تفاوت به جز تفاوت در دستگاه اندازه‌گیری به انتخاب ویژگی با ۵۰۰۰ واکسل در ROI نیز مربوط است. روش validation نیز در این پژوهش با دو پژوهش قبلی تفاوت‌هایی دارد که می‌تواند اندکی از اختلافات مشاهده شده را مسبب شود.

همچنین نتایج نشان می‌دهند ROI های جدا الگوهای فضایی نسبتاً مشابهی برای اطلاعات قطعه‌های موسیقی به دست می‌دهند و بنابراین به نظر می‌رسد اطلاعات به صورت سلسله مراتبی encode می‌شود. (همان ادعای اولیه پژوهش)

یک نکته جالب دیگر هم این است که قطعه‌هایی که اشتباه کلسیفای می‌شوند با احتمال بیشتری با قطعه دیگری از ژانر خودشان یا قطعه‌ای که از لحاظ صدایی به آن نزدیک‌تر است اشتباه می‌شوند (مثل کانتری با rocknroll) و این هم همان صحبتی است که بالاتر شد و تفاوت قطعه‌ها را به داشتن/نداشتن سازهای کوبه‌ای یا آواز انسان تقسیم کردیم.

قسمت سوم: بررسی نقاط فعال مغز در هنگام شنیدن موسیقی

۱. روش کار بدین صورت است: داده‌های کل ران‌ها خوانده شده و سپس از روی قسمت `textdata` زمان‌های تحریک‌ها بدست می‌آید. از روی همان قسمت ژانرها را نیز جدا می‌کنیم. یک خط هم کامنت شده که با اجرا شدن آن در دفعه اول فایل‌های ران را به اسکن‌ها تجزیه می‌کند تا قابل کاربری برای `spm` شوند. (۳) اسکن اول در `batch` استفاده نمی‌شوند) در ادامه نیز زمان مربوط به ژانرهای مختلف بدست آمده تا در `batch` پس از تعریف `condition` ها، زمان `onset` آن‌ها نیز داده شود. (برای هر ران این اطلاعات بدست می‌آید). در ادامه کد نیز ۲۵ کانترست اضافه شده که ۵ تای اول به جدا کردن هر کدام از ژانرها مربوط است. ۲۰ کانترست بعدی به مقایسه دو به دوی کانترست‌ها مربوط است (برای مثال یک بار `ambient` – `symphonic` و یک بار `symphonic` – `ambient`). ضمناً ما در `condition`، مشتق‌های اول این تحریک‌ها را نیز به عنوان رگرسور اضافه نموده‌ایم و بنابراین با احتساب رگرسور بایاس ۱۱ رگرسور خواهیم داشت.

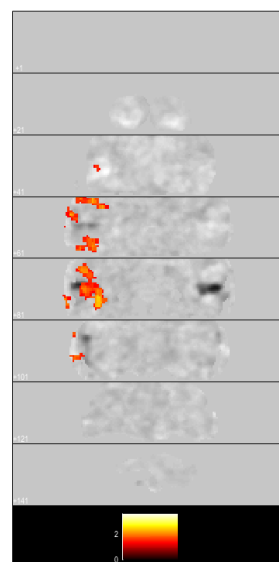
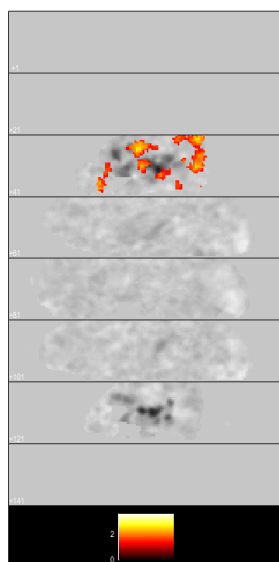
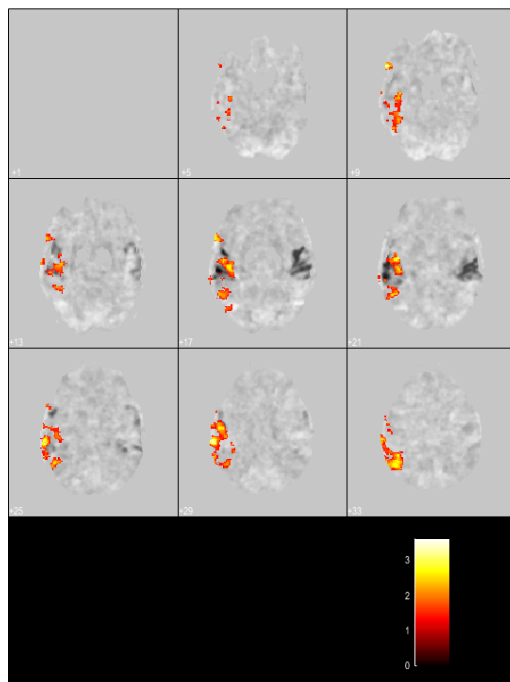
برای پاسخ به سوال اول در `spm` ابتدا با تعریف پنج `condition` برای ژانرها و دادن ۱۵۰ اسکن (۳ اسکن اول حذف شده) و زمان `onset` تحریک ژانرها در `condition` مربوطه این `batch` را ران می‌کنیم تا عملیات کانولوشن با `HRF` انجام شود. سپس ۲۵ کانترست گفته شده را اضافه می‌کنیم و این کار را با `contrast manager` انجام می‌دهیم تا برای هر ران این کارها را تکرار نکنیم! (قسمت `replicate` را فعال می‌کنیم) با اعمال این کانترست‌ها بر روی داده‌های هر ران به سادگی در فولدر خروجی که تعریف کرده ایم فایل‌های کانترست (`con_00ij`) و فایل‌های شامل `t-value` ها (`spmT_00ij`) تولید می‌شوند. با توجه به کانترست‌هایی که تعریف کردیم، از ۲۵ فایل `t-value` ابتدایی که تولید می‌شوند، ۵ فایل اول مربوط به همین قسمت هستند. از آنجایی که این فایل‌ها با نوشتن کد تولید نشده‌اند و در سوال نمایش آن‌ها نیز خواسته نشده نتایج آن‌ها قابل ارائه در این قسمت نیستند! البته در قسمت‌های بعد از آن‌ها به دفعات استفاده شده و می‌توانید از انجام عملیات این قسمت توسط خودمان اطمینان حاصل کنید!

همچنین ماتریس $z - value$ ها با دستور `spm_t2z` ساخته شدند و پارامتر ورودی درجه آزادی از شکل های `results` در `spm` که درجه آزادی متغیر `t` را نشان می‌دادند برابر ۱۳۵ انتخاب شد. این ماتریس ها نیز قابل ارائه نیستند و به `screenshot` ای از یک ران که حاوی فایل های `contrast` و `t - value` و `z - value` باشد بسنده می‌کنیم! چند خط کد در فایل اصلی کامنت شده است که وظیفه بدست آوردن `z - value` ها را داشته و با توجه به چیدمان فایل ها در سیستم ما نوشته شده است و در صورت تفاوت داشتن با سیستم اجرایی شما، اجرای آن‌ها فایده‌ای نخواهد داشت.

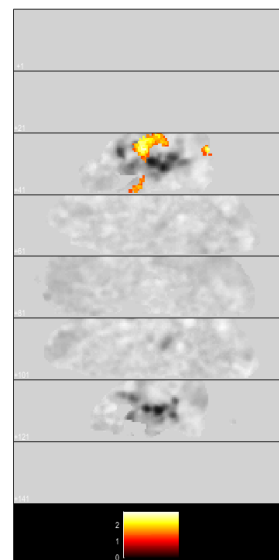
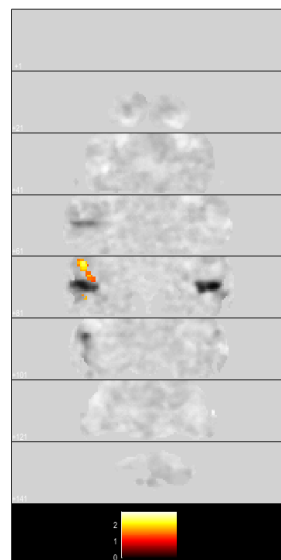
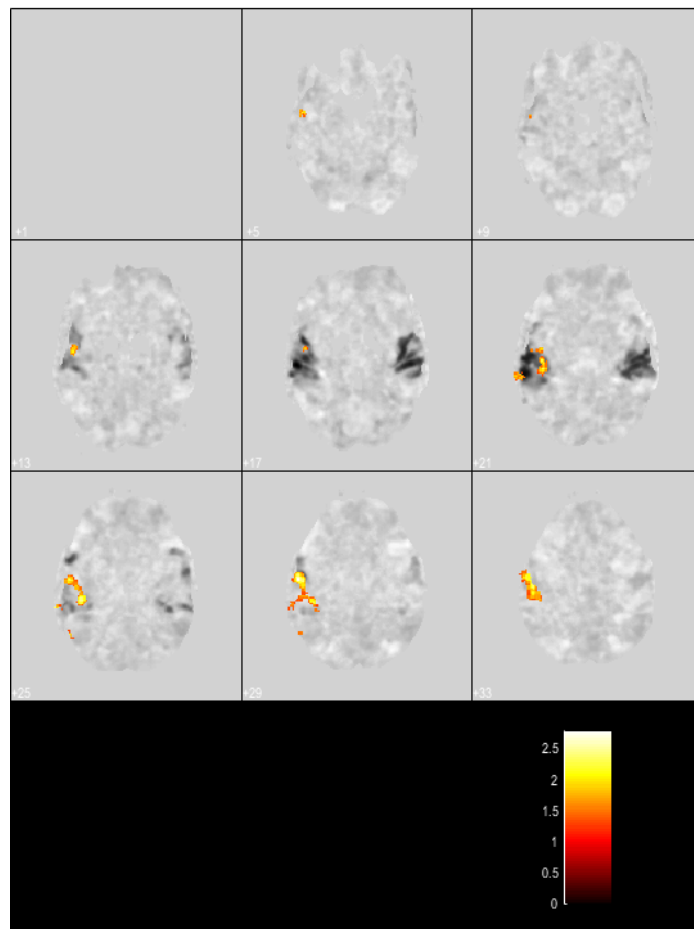
beta_0001.nii	con_0019.nii	con_0048.nii	spmT_0019.nii	spmT_0048.nii	spmZ_0014.hdr
beta_0002.nii	con_0020.nii	con_0049.nii	spmT_0020.nii	spmT_0049.nii	spmZ_0014.img
beta_0003.nii	con_0021.nii	con_0050.nii	spmT_0021.nii	spmT_0050.nii	spmZ_0015.hdr
beta_0004.nii	con_0022.nii	cont_batch.mat	spmT_0022.nii	spmZ_0001.hdr	spmZ_0015.img
beta_0005.nii	con_0023.nii	cont_batch2.mat	spmT_0023.nii	spmZ_0001.img	spmZ_0016.hdr
beta_0006.nii	con_0024.nii	contrasts_batch.mat	spmT_0024.nii	spmZ_0002.hdr	spmZ_0016.img
beta_0007.nii	con_0025.nii	mask.nii	spmT_0025.nii	spmZ_0002.img	spmZ_0017.hdr
beta_0008.nii	con_0026.nii	outputs	spmT_0026.nii	spmZ_0003.hdr	spmZ_0017.img
beta_0009.nii	con_0027.nii	ResMS.nii	spmT_0027.nii	spmZ_0003.img	spmZ_0018.hdr
beta_0010.nii	con_0028.nii	RPV.nii	spmT_0028.nii	spmZ_0004.hdr	spmZ_0018.img
beta_0011.nii	con_0029.nii	SPM.mat	spmT_0029.nii	spmZ_0004.img	spmZ_0019.hdr
con_0001.nii	con_0030.nii	spmT_0001.nii	spmT_0030.nii	spmZ_0005.hdr	spmZ_0019.img
con_0002.nii	con_0031.nii	spmT_0002.nii	spmT_0031.nii	spmZ_0005.img	spmZ_0020.hdr
con_0003.nii	con_0032.nii	spmT_0003.nii	spmT_0032.nii	spmZ_0006.hdr	spmZ_0020.img
con_0004.nii	con_0033.nii	spmT_0004.nii	spmT_0033.nii	spmZ_0006.img	spmZ_0021.hdr
con_0005.nii	con_0034.nii	spmT_0005.nii	spmT_0034.nii	spmZ_0007.hdr	spmZ_0021.img
con_0006.nii	con_0035.nii	spmT_0006.nii	spmT_0035.nii	spmZ_0007.img	spmZ_0022.hdr
con_0007.nii	con_0036.nii	spmT_0007.nii	spmT_0036.nii	spmZ_0008.hdr	spmZ_0022.img
con_0008.nii	con_0037.nii	spmT_0008.nii	spmT_0037.nii	spmZ_0008.img	spmZ_0023.hdr
con_0009.nii	con_0038.nii	spmT_0009.nii	spmT_0038.nii	spmZ_0009.hdr	spmZ_0023.img
con_0010.nii	con_0039.nii	spmT_0010.nii	spmT_0039.nii	spmZ_0009.img	spmZ_0024.hdr
con_0011.nii	con_0040.nii	spmT_0011.nii	spmT_0040.nii	spmZ_0010.hdr	spmZ_0024.img
con_0012.nii	con_0041.nii	spmT_0012.nii	spmT_0041.nii	spmZ_0010.img	spmZ_0025.hdr
con_0013.nii	con_0042.nii	spmT_0013.nii	spmT_0042.nii	spmZ_0011.hdr	spmZ_0025.img
con_0014.nii	con_0043.nii	spmT_0014.nii	spmT_0043.nii	spmZ_0011.img	
con_0015.nii	con_0044.nii	spmT_0015.nii	spmT_0044.nii	spmZ_0012.hdr	
con_0016.nii	con_0045.nii	spmT_0016.nii	spmT_0045.nii	spmZ_0012.img	
con_0017.nii	con_0046.nii	spmT_0017.nii	spmT_0046.nii	spmZ_0013.hdr	
con_0018.nii	con_0047.nii	spmT_0018.nii	spmT_0047.nii	spmZ_0013.img	

۲. برای پاسخ به سوال ۲ و ۳، مقادیر $z - value$ متناظر با هر کانترست را برای تمامی ران ها بدست آورده و میانگین می‌گیریم. اشکال مشاهده شده مربوط به میانگین‌ها هستند. نتیجه را در انتهای تصاویر آورده‌ایم.

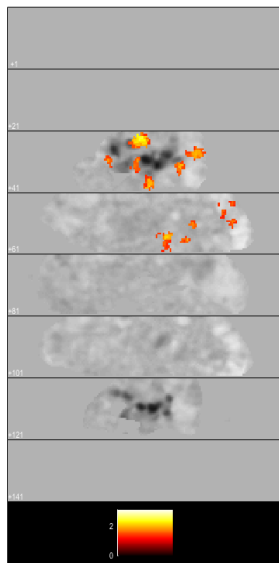
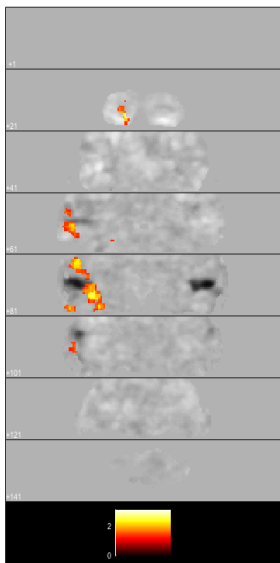
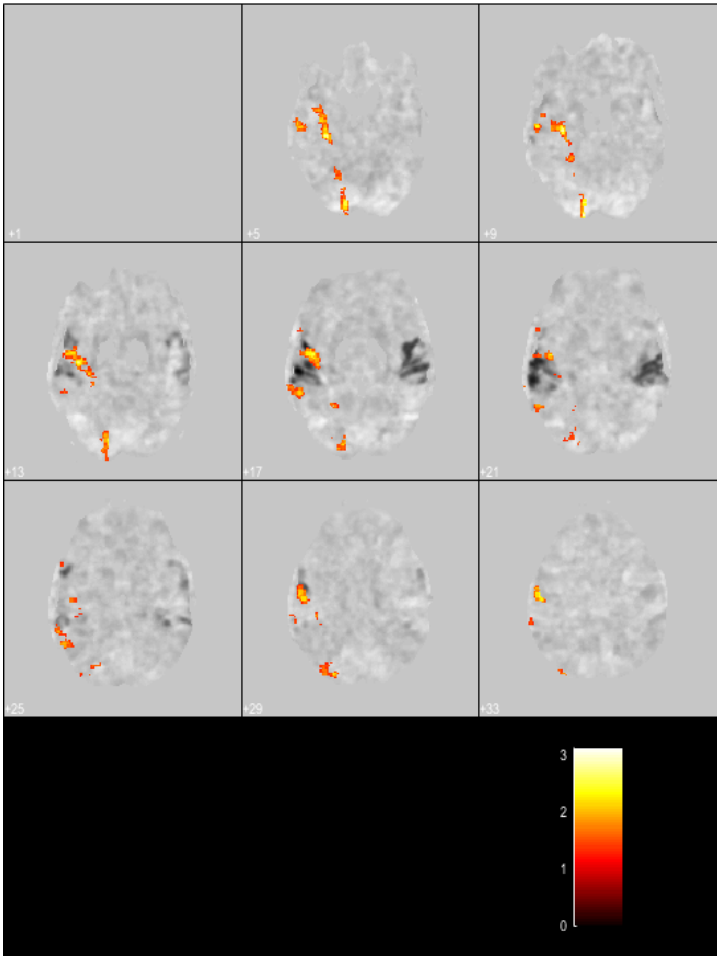
Ambient



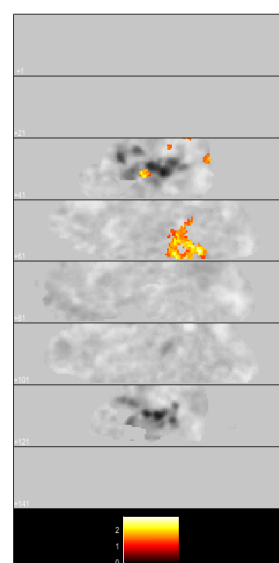
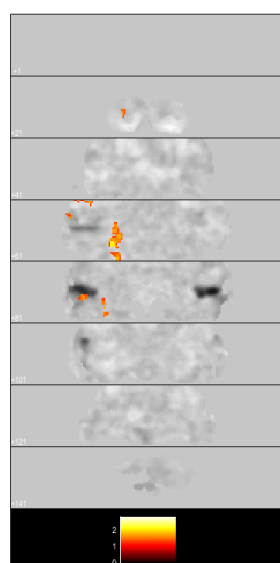
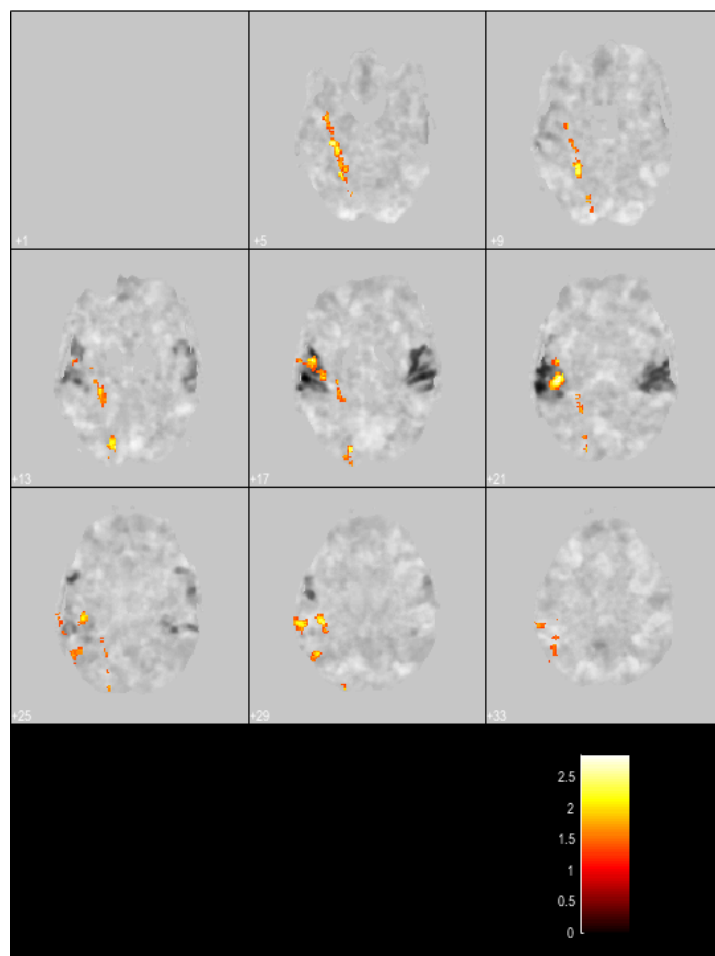
RocknRoll



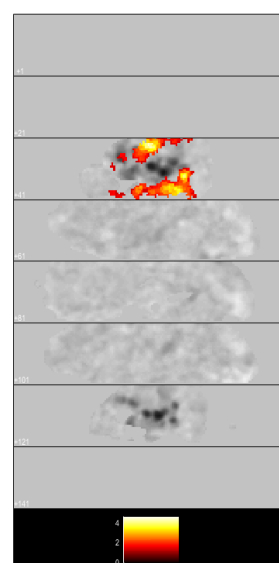
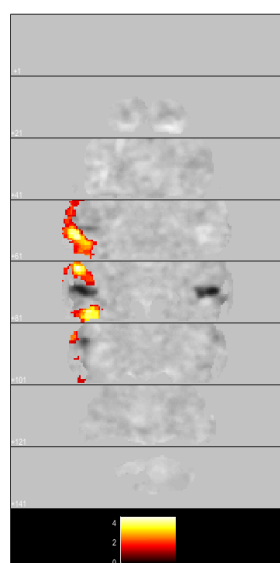
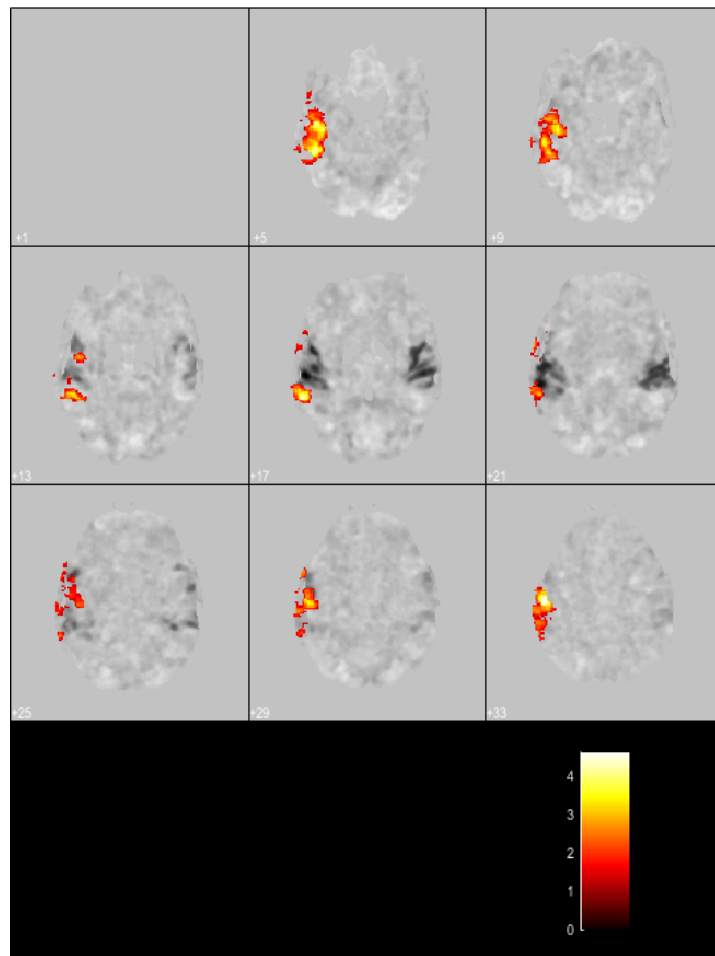
Country



Metal

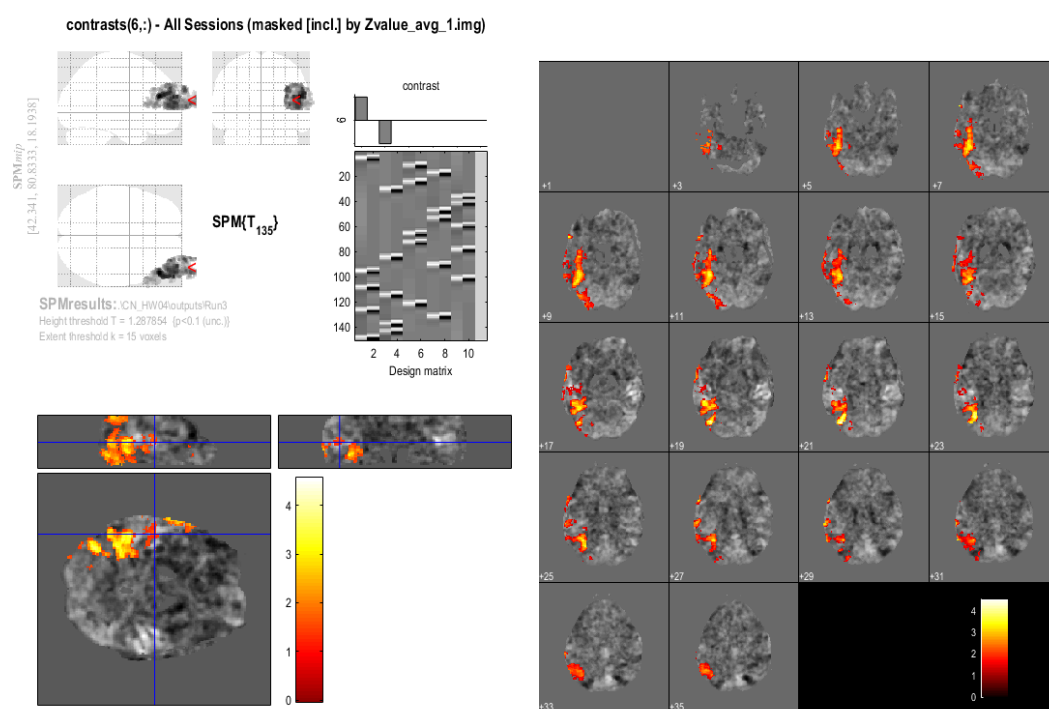


Symphonic

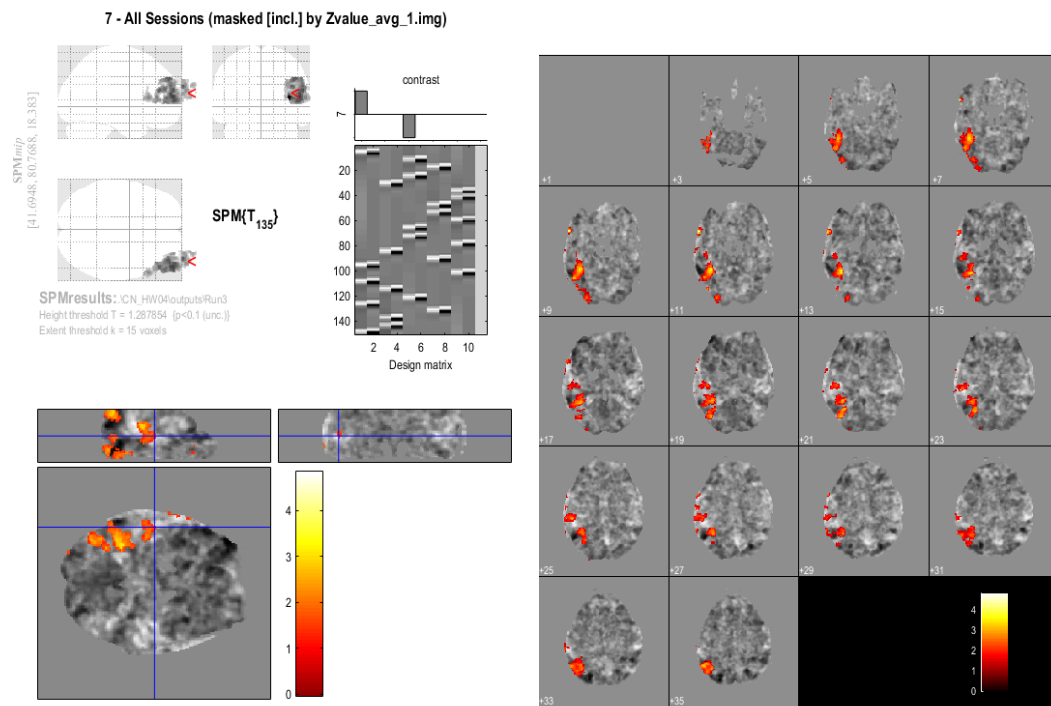


۳. از آنجایی که تصویر از بالا، تقریباً تمامی اطلاعات را دارد، برای این حالت‌های مقایسه دو به دو فقط این تصویر را می‌آوریم و از آوردن دو زاویه دیگر خودداری می‌کنیم. هر دو مجموعه متوالی در تصاویر زیر، مربوط به کانترست‌های قرینه هم هستند. (نتیجه در انتهای تصاویر)

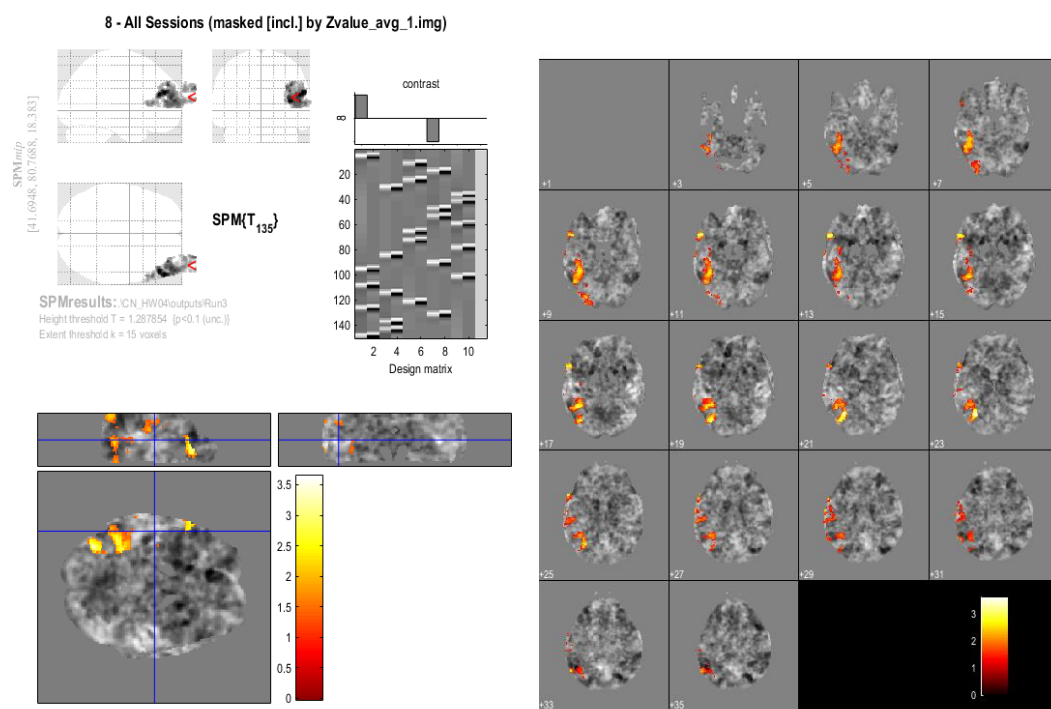
Ambient – RocknRoll



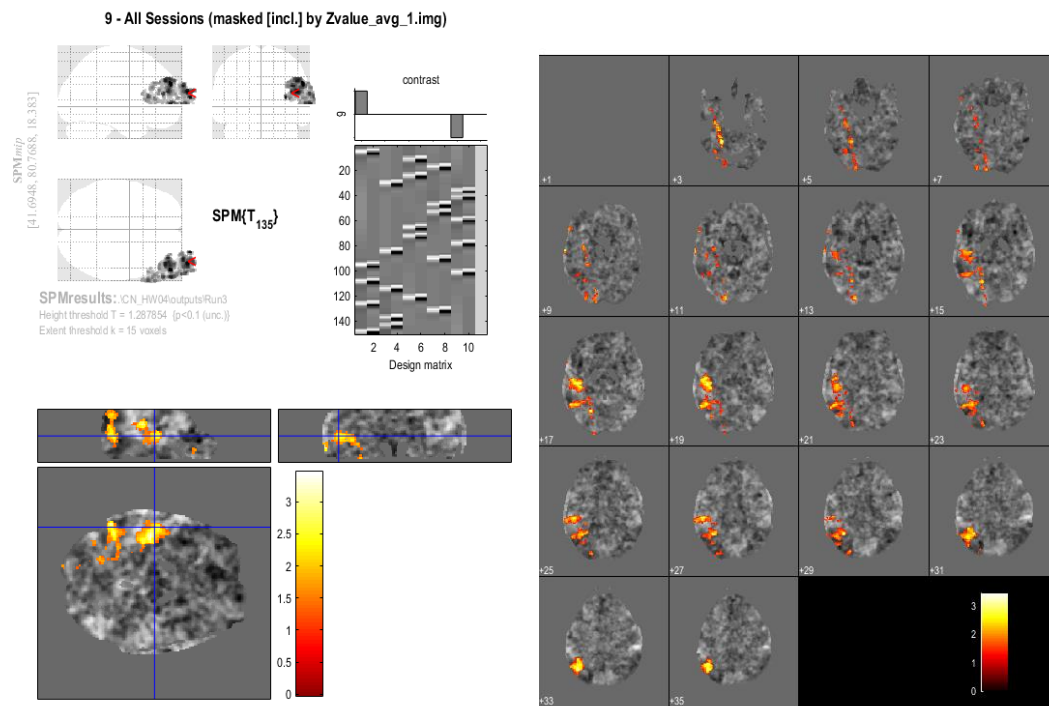
Ambient – Country



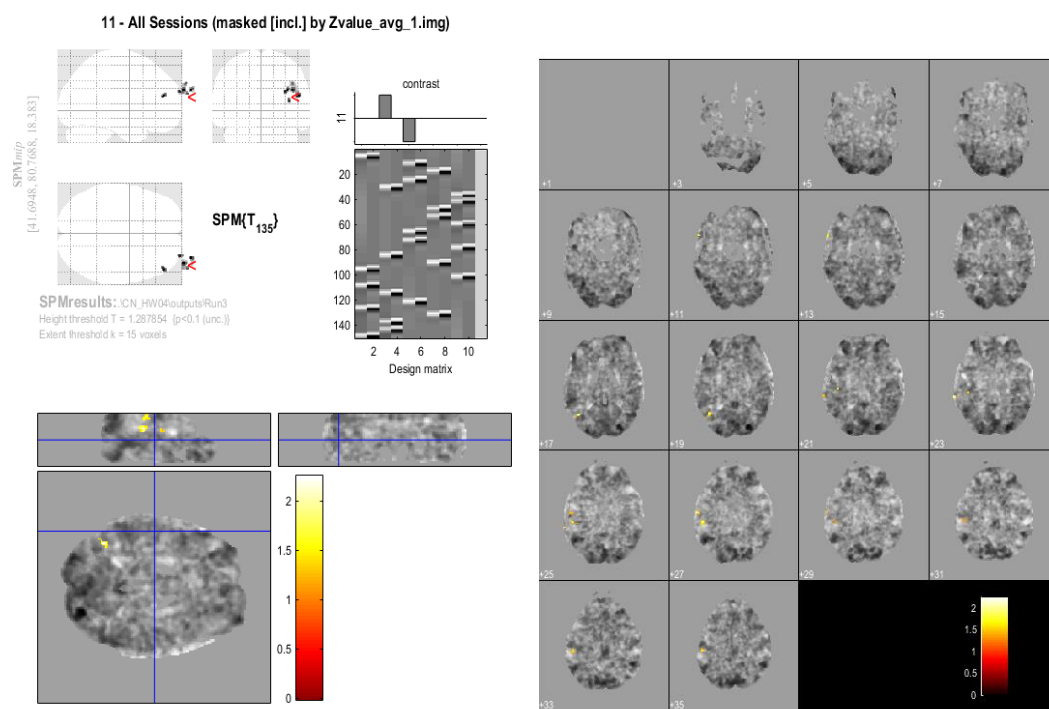
Ambient – Metal



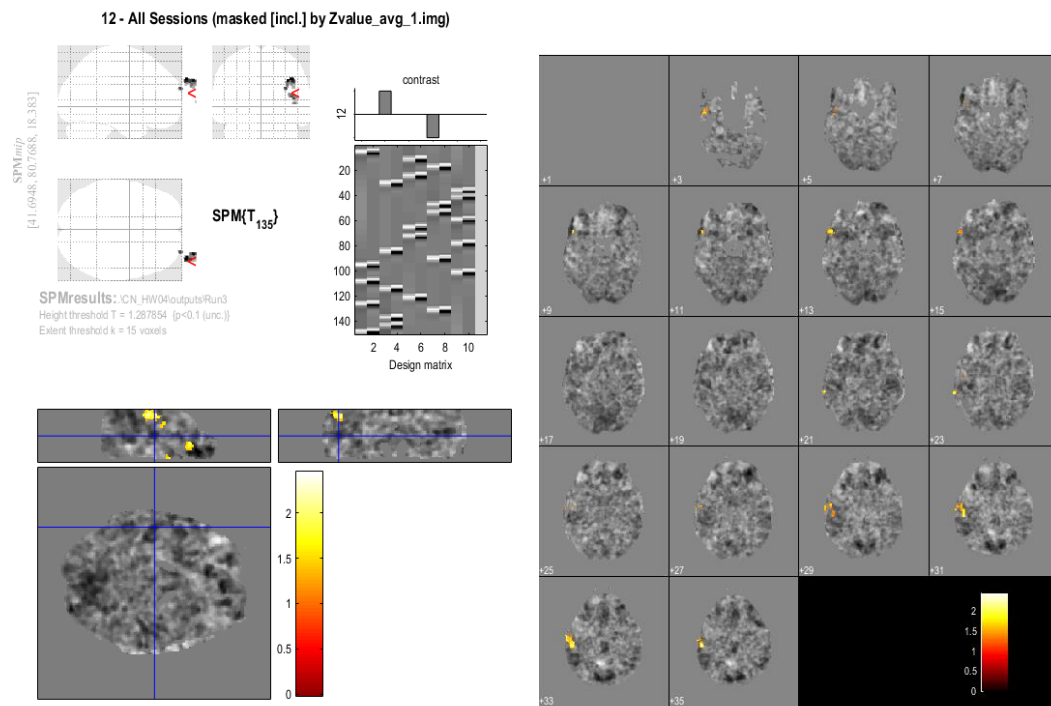
Ambient – Symphonic



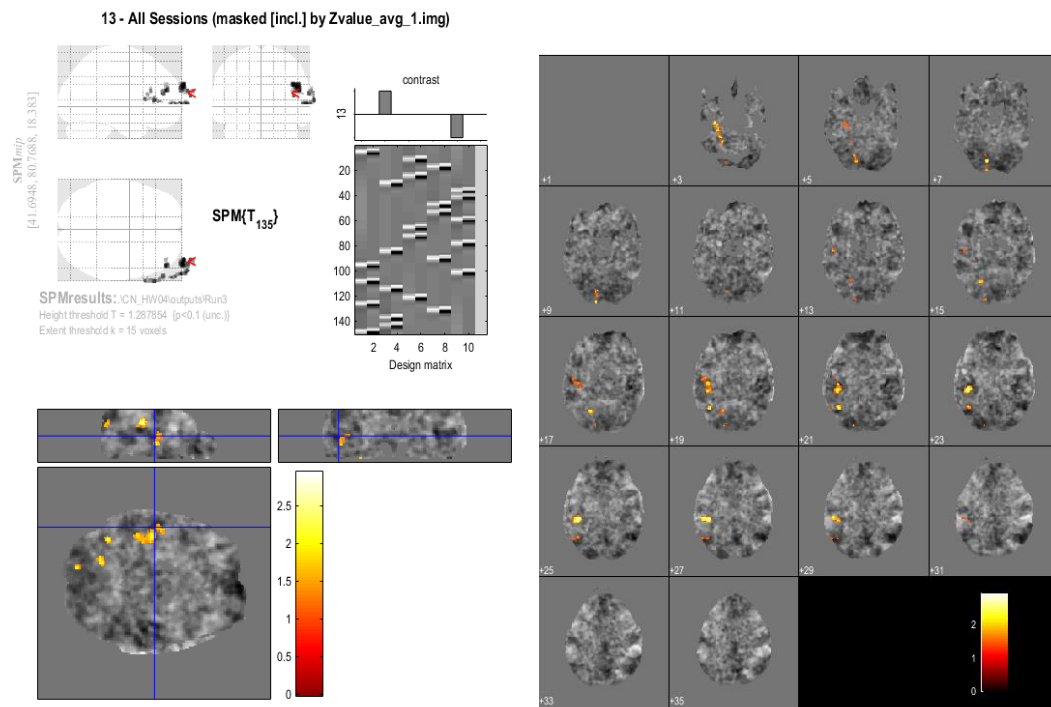
RocknRoll – Country



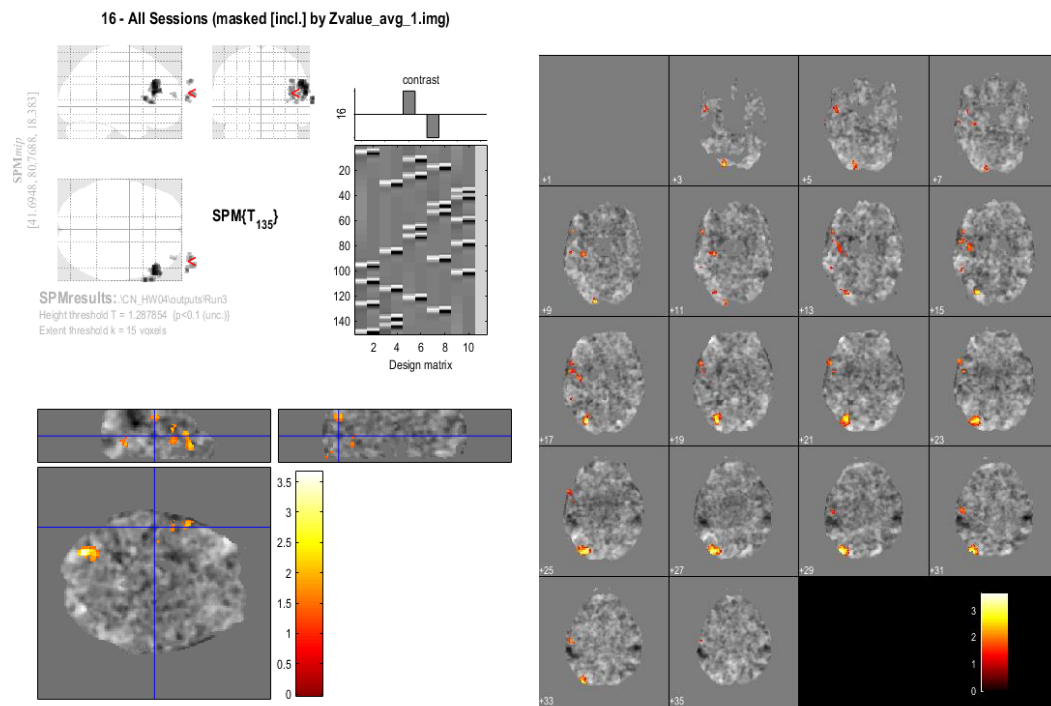
RocknRoll – Metal



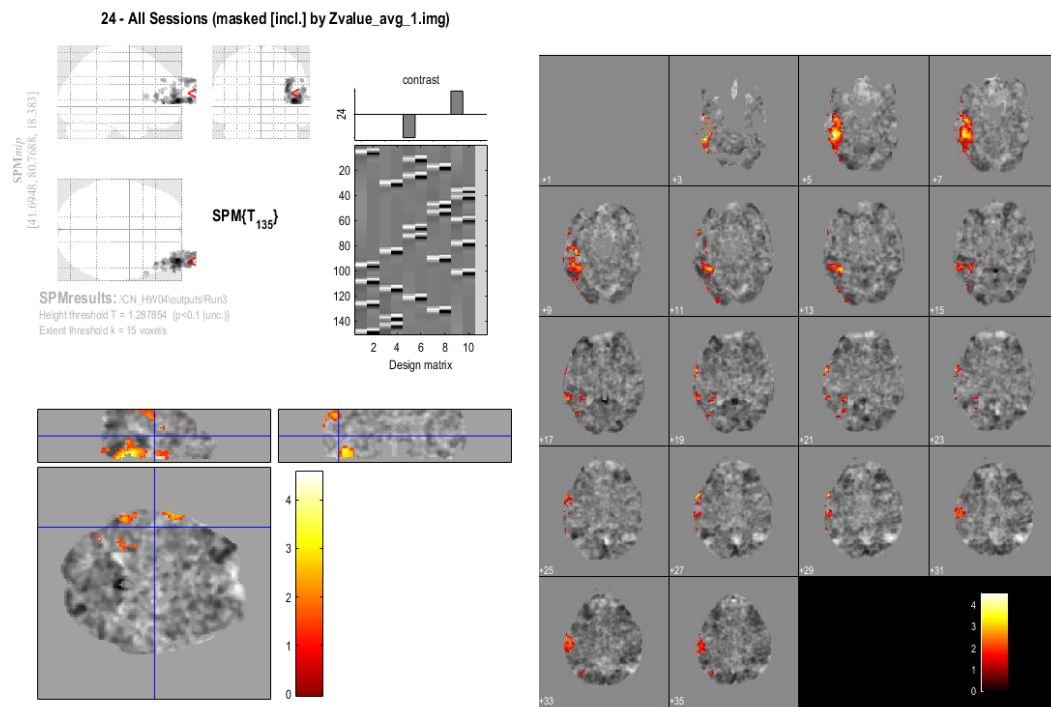
RocknRoll – Symphonic



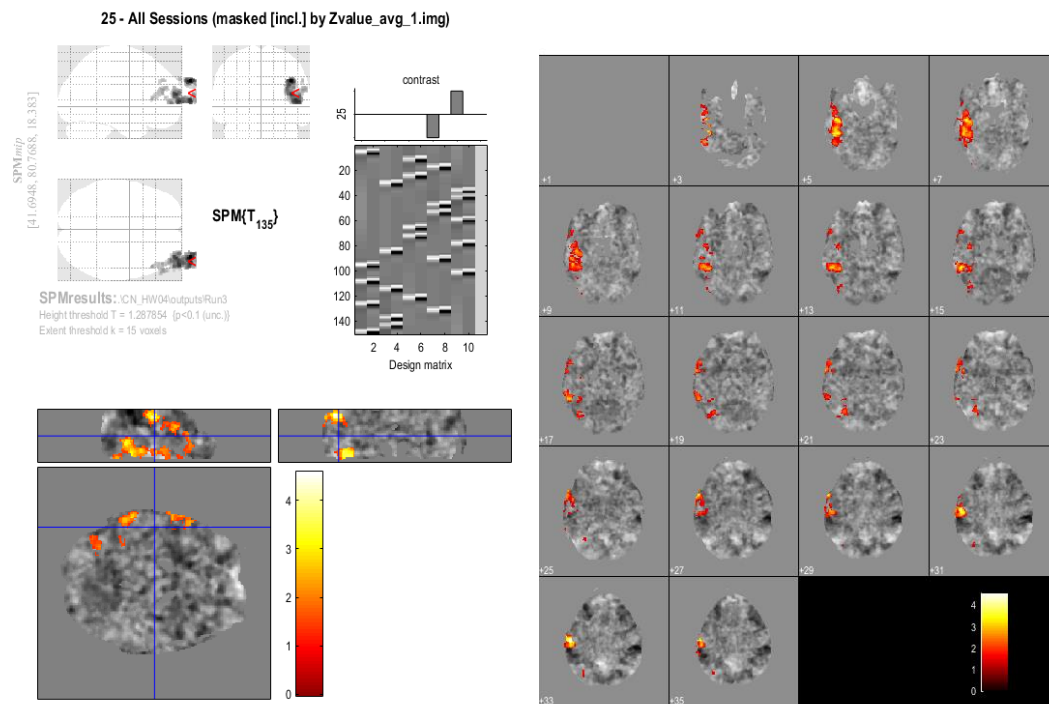
Country – Metal



Symphonic – Country



Symphonic – Metal



نتیجه سوال ۲ و ۳:

سوال ۲: در هر ۵ ژانر می‌توان مشاهده کرد که ناحیه رنگی همان ناحیه auditory cortex است و با تصاویر موجود در اینترنت سازگاری دارد. در بعضی از ژانرها مسیر اطلاعات تا این ناحیه نیز فعال مشاهده می‌شود ولی الگوی کلی تصاویر تفاوت چندانی ندارند و مطابق انتظار ناحیه شنوایی سر را فعال کرده‌اند.

فعال شدن این ناحیه در حالت *symphonic* و *rocknroll* شدیدتر از بقیه حالت‌ها است و در دو حالت *country* و *metal* مسیر خطی نیز فعال به نظر می‌رسد.

سوال ۳: در شکل ابتدایی این سوال می‌توان ملاحظه کرد که *ambient* با *rocknroll* تفاوت معناداری دارند و ناحیه شنوایی به اختلاف این دو حساس است (برای این ناحیه این دو ژانر واقعاً تفاوت دارند). (جدول زیر تفاوت‌ها را به صورت کیفی با بررسی چشمی بدست می‌دهد!)

تفاوت	Ambient	Rocknroll	Country	Metal	Symphonic
Ambient		زیاد	کم	زیاد	زیاد(!)
Rocknroll	زیاد		کم	کم	متوسط
Country	کم	کم		متوسط	زیاد
Metal	زیاد	کم	متوسط		زیاد
Symphonic	زیاد(!)	متوسط	زیاد	زیاد	

نقاط فعال مغز همان نقاط سوال ۲ هستند که باید هم باشند چون تحریک‌های صوتی در آن ناحیه پاسخ داده می‌شوند ولی می‌توان میزان تمایز مغز و ناحیه شنیداری برای ژانرهای مختلف را با فعالیت تفضلی که با کانترست مناسب تعریف شده مشاهده کرد. بر این اساس انتظار داریم پس از کلسیفای کردن نیز آن‌هایی که در جدول بالا تفاوت زیادی داشتند راحت‌تر کلسیفای شده ولی مواردی که تفاوت کم داشتند احتمال دارد misclassify بشوند. همچنین جالب است که مشاهده کنیم ارتباط «صدایی» یا «هم‌صدایی» که مطرح شد به طور کلی در اینجا نیز قابل مشاهده است و امکان دسته‌بندی معنادارتری از ژانرها را فراهم می‌کند. (یا دسته‌بندی در سطح بالاتر)

۴. آن‌چه مشاهده شد فعال شدن ناحیه شنیداری و مسیری خط مانند بود که در اینترنت هم پس از جستجو به نظر می‌رسد به عملیات شنیدن مربوط باشد. ناحیه‌های دیگر نیز کم و بیش فعال هستند از جمله پیشانی و پشت سر. ولی فعالیت معنادار مطابق آن‌چه در قسمت‌های قبل گفته شد در یک سمت مغز بوده و به ناحیه شنیداری مربوط است و نتایج با واقعیت سازگار است. یک مورد مهم دیگر عملیات توجه کردن است که در بعضی زمان‌ها مورد اهمیت بوده (مثل جلب شدن توجه به موردی در قطعه موسیقی یا در پاسخ دادن به سوال catch) که این باعث فعال شدن سایر قسمت‌های مشاهده شده است.

قسمت چهارم: طبقه‌بندی ژانرهای موسیقی متفاوت

۱. برای تعریف design matrix در این قسمت، برای هر run متناظر با هر تحریک یک رگرسور در نظر می‌گیریم که قبل از کانولوشن با hrf از زمان onset آن به مدت ۶ ثانیه ۱ و در بقیه جاها صفر می‌باشد. سایر تنظیمات در 1st – level analysis ، batch ، نمونه‌ای که به پیوست قرار دارد قابل ملاحظه است (run – 01_q4_batch.mat) نهایتاً به ازای هر تحریک t(z) value ای در فضای ۹۲۱۶۰۰ واکسلی بدست می‌آوریم.

۲. از تابع anova1 استفاده می‌کنیم. اکنون به ازای هر واکسل یک p-value بدست آورده‌ایم که نشان می‌دهد آن پاسخ آن واکسل چقدر می‌تواند بین تحریک‌های مختلف تمایز دهنده باشد. با فرض مقدار آستانه‌ای واکسل‌های تمایزدهنده (به تعبیری informative) را انتخاب می‌کنیم (واکسل‌هایی با کوچکترین p-value ها).

۳. با این که در سوال قبل ابعاد فضا (تعداد feature ها) را کاهش دادیم اما ممکن است همچنان تعداد این ابعاد (تعداد واکسل‌ها) از تعداد مشاهدات (تعداد تحریک‌ها) بیشتر باشد که بی‌معنی است. بنابراین در صورت رخداد چنین حالتی با PCA ابرصفحه‌ای که داده‌ها در آن قرار دارند را پیدا می‌کنیم و داده‌های تصویر شده به این ابرصفحه را به کلاسیفایرمان می‌دهیم. البته در اینجا بیشتر از 24 بعد نگه نمی‌داریم که توضیحات بیشتر در بخش پنجم داده شده است.

از تابع fitcdiscr برای کلاسیفای کردن استفاده می‌کنیم که قابلیت طبقه‌بندی با multiple classes را هم دارد. اگر از threshold > 5e-4 برای p-value های خروجی anova1 و همان داده‌های training برای بررسی دقت طبقه‌بندی استفاده کنیم، بدست می‌آید: (راجع به انتخاب threshold در بخش‌های بعدی توضیح داده شده است)

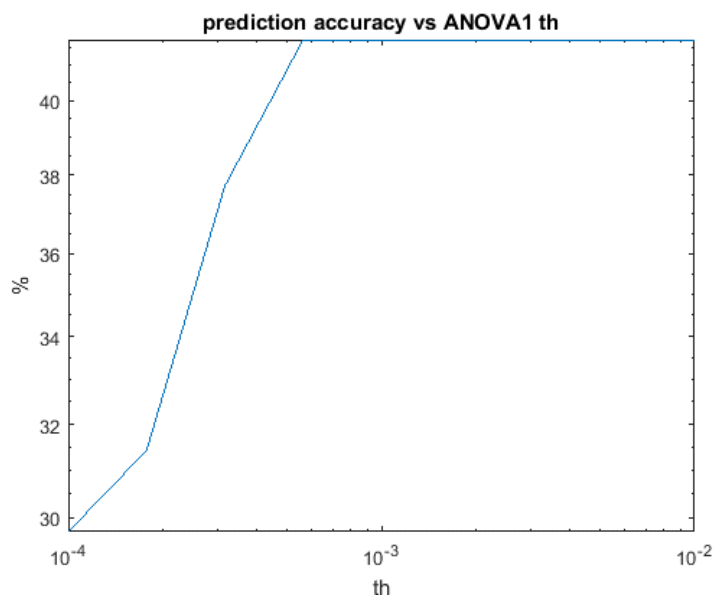
```
percent of correct predictions using fitcdiscr (p_val th = 0.006): 41.714%
```

۴. در هر اجرا، دیتاهای یک run را کنار گذاشته، با اجرای anova1 ، p-value های متناظر با هر واکسل را انتخاب می‌کنیم. ترجیح می‌دهیم به جای اعمال threshold بر این p-value ها، تعداد واکسل‌های تمایزدهنده را ثابت نگه داریم. لذا تعداد مشخصی (با توجه به توضیحات سوال 5، حدود 15 هزار) از واکسل‌ها را انتخاب می‌کنیم و ابعاد فضا را برای رسیدن به زیر 25 (left out run) ، 25 تا تحریک دارد) با PCA کاهش می‌دهیم. البته برای دوری از high variance، صرفاً 20 بعد را نگه می‌داریم. نهایتاً کلاسیفیکیشن را انجام داده و طبقه‌بند حاصل را روی run ای که کنار گذاشته بودیم بررسی می‌کنیم. دقت میانگین

کلاسفایرها برای run های مختلف را میزانی از عملکرد کلاسفرمان می گیریم. اگر خطای این عملکرد را با $\frac{\sigma}{\sqrt{7-1}}$ محاسبه کنیم داریم:

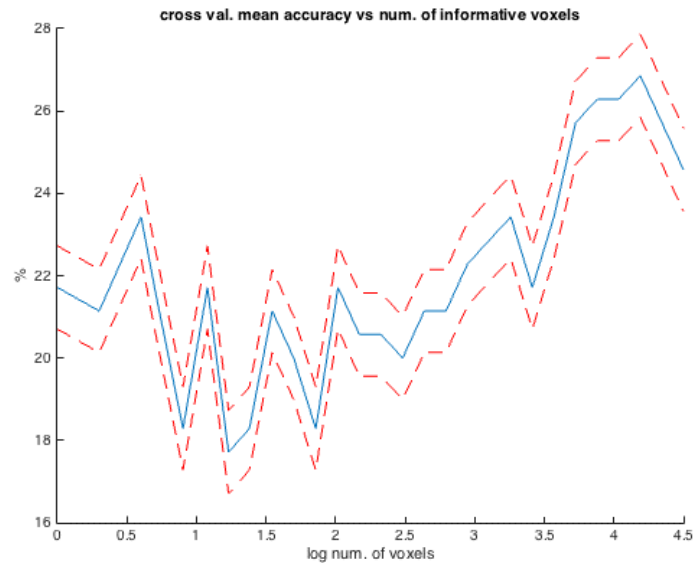
max accuracy: 26.86% , error bound: 1.01 for num. of voxels: 15475

۵. در تکرار سوال سوم به ازای threshold ها مختلف نمودار زیر بدست می آید:



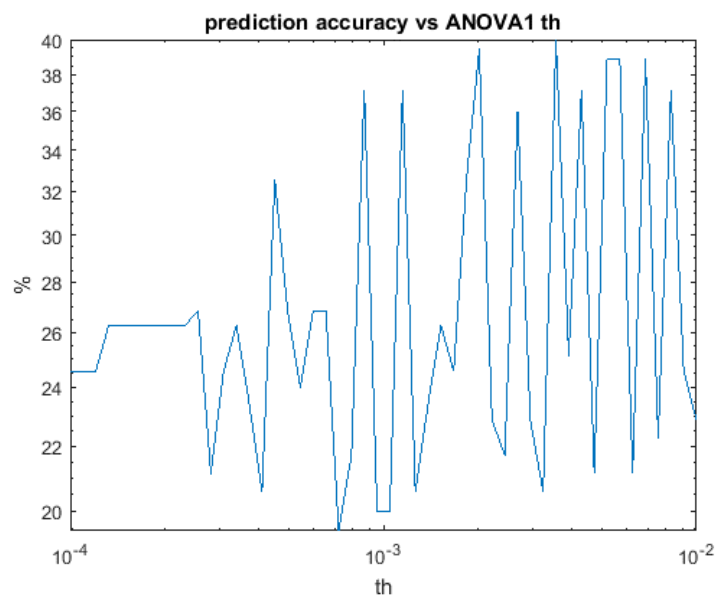
همانطور که ملاحظه می گردد استفاده از مقادیر آستانه بزرگ و اعمال PCA عملکرد بهتری داشته است. بازهم برای دوری از مسئله high variance کوچکترین th را ترجیح می دهیم (حدود 5e-4). در این حالت حدود 16 واکسل تمایزدهنده انتخاب می شوند.

کار مشابهی را برای انتخاب تعداد واکسل ها در سوال 4م انجام داده ایم. نتیجه در شکل زیر قابل ملاحظه است. برای رسیدن به دقت حداکثری حدود 15000 واکسل را انتخاب می کنیم و فضا را با PCA به 20 بعد کاهش می دهیم.

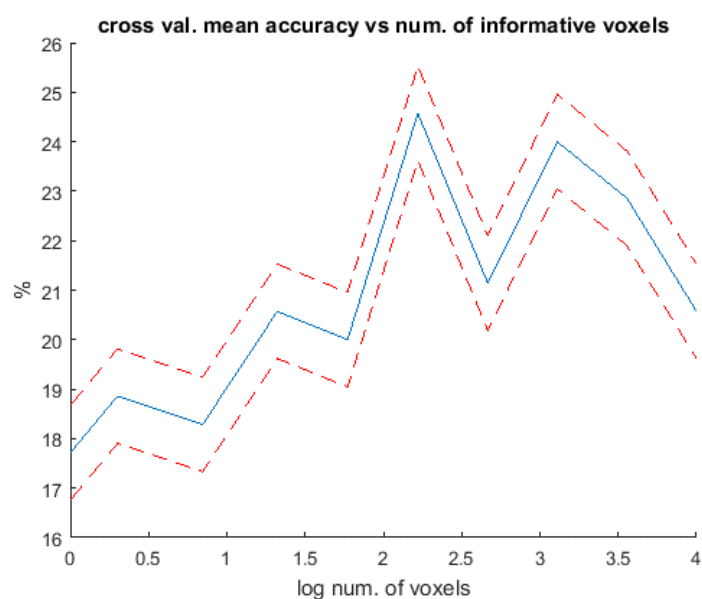


۶. به نظر می‌رسد در همه این روش‌ها تلاش می‌شود **famaliywise error rate** (همان **false positive**، در اینجا یعنی فرض **null** را رد کنیم (مثلاً بگوییم واکسلی تمایزدهنده است) در حالی که این گونه نبوده). مستقل از اندازه ابعاد فضا محدود بماند. دو روش شاخص آن همانطور که در صورت سوال اشاره شده است **Benforreni correction** و **BH algorithm** است. در روش اول با کوچک کردن مقدار آستانه **p-value** (متناسب با معکوس ابعاد فضا) این کار انجام می‌گیرد. مثلاً بدین صورت که اگر سابقاً احتمال یک **false positive** برای یک واکسل **a** بوده است، با انتخاب آستانه جدید **a/m** (که **m** ابعاد فضا است) احتمال یک **false positive** برای **m** واکسل کمتر از **a** می‌ماند. در روش دوم مستقل از اندازه **p-value** ها آستانه را انتخاب می‌کنیم بدین صورت که ابعاد را برحسب **p-value** هایشان مرتب می‌کنیم و با توجه به نرخ خطای مدنظرمان تعدادیشان را انتخاب می‌کنیم.

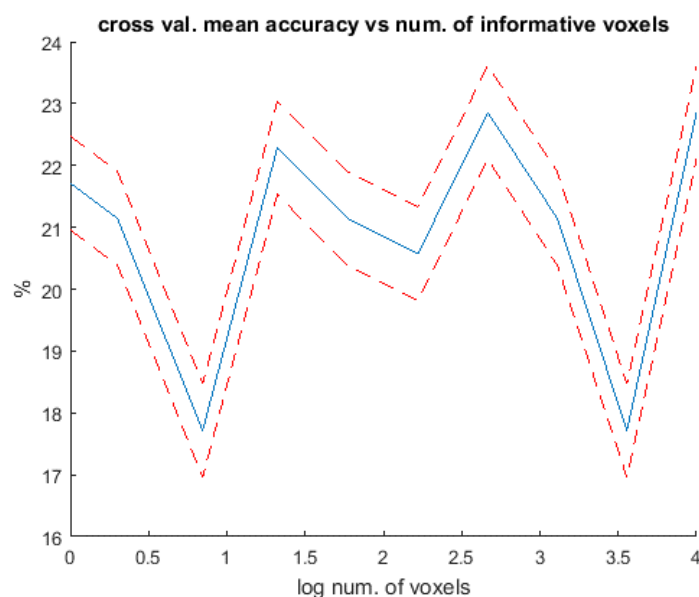
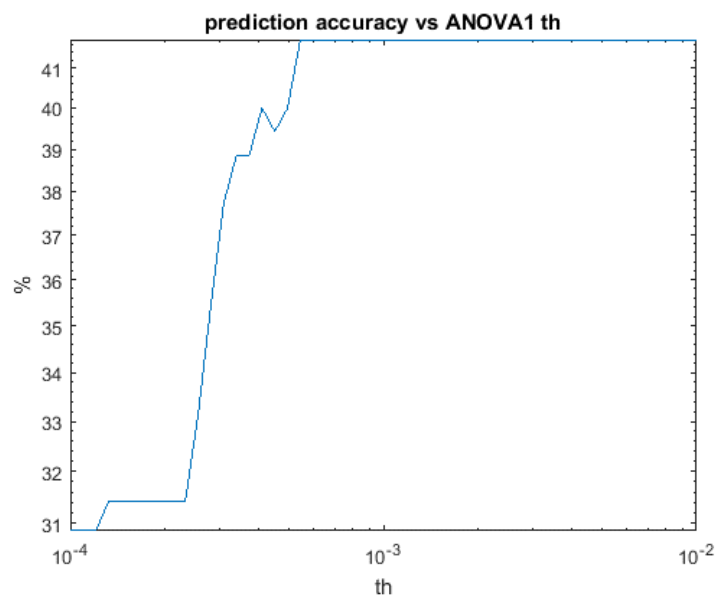
۷. برای کلاسیفایر **SVM**، دقت طبقه بندی روی دیتاهای **training** به عنوان تابعی از **threshold** در **anova1** در ادامه رسم شده است. نکته عجیب نوسانات بسیار سریع دقت با **threshold** است که انتخاب آستانه را دشوار می‌کند اما اندازه این حداکثر دقت حدود 40 درصد است.



مشابه قبل با cross validation و تغییر تعداد واکسل‌ها انتخابی، نمودار زیر حاصل شد. به نظر می‌رسد عملکرد SVM در مجموع ضعیف‌تر از LDA است.



مراحل فوق را برای کلاسیفایری که از logistic regression استفاده می‌کند تکرار کردیم. نتایج در ادامه ارائه شده است. به نظر می‌رسد عملکرد این کلاسیفایر نیز مطلوب نیست.



۸. ماتریس confusion در زیر رسم شده است. از نظر این که عمده وزن‌ها روی قطر اصلی متمرکز باشد نتایج به مقاله (و انتظار ما!) نزدیک است. اما از نظر این که ambient بهتر از بقیه تشخیص داده شوند و این که metal و rocknroll اشتباها به جای هم تشخیص داده شوند شاید نتایج مطابقت نداشته باشد. Behavioural confusion عملکرد انسان (با پرسش و پاسخ) در تشخیص ژانر موسیقی را نشان می‌دهد. به نظر نمی‌رسد شباهت معناداری بین نتایج ما و ساختار Behavioural confusion به جز در نزدیک به قطری بودنش وجود داشته باشد.

	Predicted					
		Ambient	Country	Metal	Rocknroll	Symphonic
real	Ambient	6	7	10	4	8
	Country	7	6	5	10	7
	Metal	2	7	16	3	7
	Rocknroll	4	9	10	7	5
	Symphonic	6	8	5	4	12

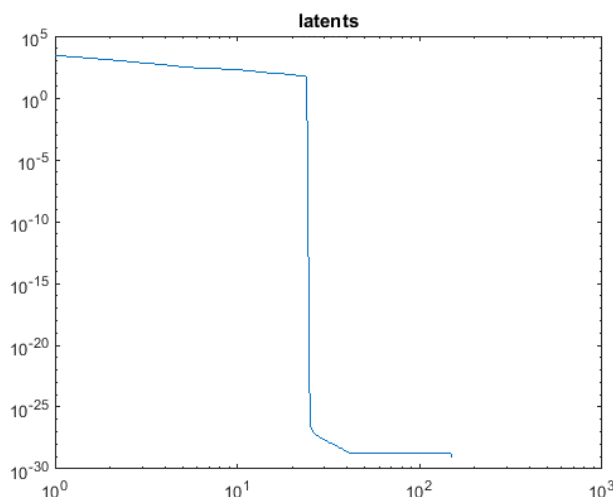
۹. خروجی به پیوست قرار دارد. اعداد 1 تا 5 خروجی به ترتیب متناظر هستند با:

Ambient, country, metal, rocknroll, symphonic

۱۰. در این بخش برخلاف بخش قبل در ابتدا به هر تحریک مستقل از ژانر آن نگاه کردیم و سعی کردیم با نسبت دادن رگرسوری به آن، برای هر تحریک یک پاسخ بیابیم. با این کار تعداد داده‌های خود را نسبت به وقتی که رگرسورها را به ژانرها نسبت می‌دادیم نیز افزایش دادیم. در این مرحله تعداد فیچرهای ما به تعداد کل واکسل‌ها و بسیار زیاد بود. لذا برای کاهش تعداد فیچرها در گام اول واکسل‌های تمایزدهنده را شناسایی کردیم (anova1) و در گام بعد در صورت لزوم ابعاد فضا را با PCA به زیر 24 بعد رساندیم. نهایتاً داده‌ها را با کلاسیفایرهای مختلف طبقه‌بندی نموده و عملکرد آن‌ها را با cross validation سنجیدیم. به نظر می‌رسد مسئله ما یک مسئله high variance است که برای بهبود عملکرد آن یا باید کلاسیفایرهای ساده‌تری انتخاب کرد (که ساده‌تر از خطی نمیشناسیم!) یا داده‌های بیشتری جمع‌آوری کرد که جواب بسیار بهتر مقاله اصلی احتمالاً به همین خاطر است.

قسمت پنجم: سوال دلخواه: (شاید وقتی دگر!)

یک مشاهده قابل تامل این است که بعد از اعمال PCA بر داده‌ای که با thresholding روی p-value های anova1 بدست آمده، تقریباً مستقل از آستانه انتخابی تقریباً 24 مقدار ویژه بزرگ وجود دارد!



این مسئله مورد قابل توجهی نبود اگر ما صرفاً یک run با 25 تحریک داشتیم و در این صورت پاسخ به هر تحریک می‌توانست یکی از آن 24 جهت بامعنی باشد. لذا از مشاهده فوق دو حدس می‌توان زد:

- تحریک‌ها در run های مختلف ثابت بوده و صرفاً جایگشتشان تغییر می‌کرده.
- با فرض درستی حدس قبل، پاسخ مغز به هر تحریک تقریباً مستقل از زمان اعمال آن تحریک و تحریک‌های دیده شده قبلی است.

اگر ادعاهای فوق درست باشد علی‌الاصول باید بتوان تحریک‌های تکراری در run های مختلف را با clustering پاسخ‌ها بدست آورد. بدین منظور با اعمال یک خوشه‌بندی ساده (kmeans) با 25 خوشه روی پاسخ دیده شده از هر واکسل (به طول 150، یک run جدا شده است) مشاهده کردیم خوشه‌هایی با اندازه دقیقاً یکسان (6) بدست آمد که اگر به genre متناظرشان نگاه می‌کردیم تقریباً یک genre غالب در هر خوشه وجود می‌داشت. لذا در ادامه تصمیم گرفتیم classifying را به جای ژانرها روی شماره تحریک‌ها (شماره متناظر با خوشه‌شان) انجام دهیم (همه این بخش با فرض این است که نمی‌دانیم چه آهنگی به عنوان تحریک استفاده شده و صرفاً حدس می‌زنیم آهنگ در run های مختلف تکرار شده پس شماره تقریبی هر آهنگ را از clustering بدست آوردیم). سپس از طبقه‌بند حاصل برای پیش‌بینی left out run استفاده می‌کنیم. نهایتاً ژانر آن تحریک را ژانر غالب خوشه‌ای که شماره آهنگ پیش‌بینی شده در آن قرار دارد در نظر می‌گیریم. نتیجه بهترین انتخاب تعداد واکسل‌ها برای cross validation عبارت است از:

`max accuracy: 25.71% , error bound: 1.09% for num. of voxels: 2395`

که نتیجه قابل توجهی است.