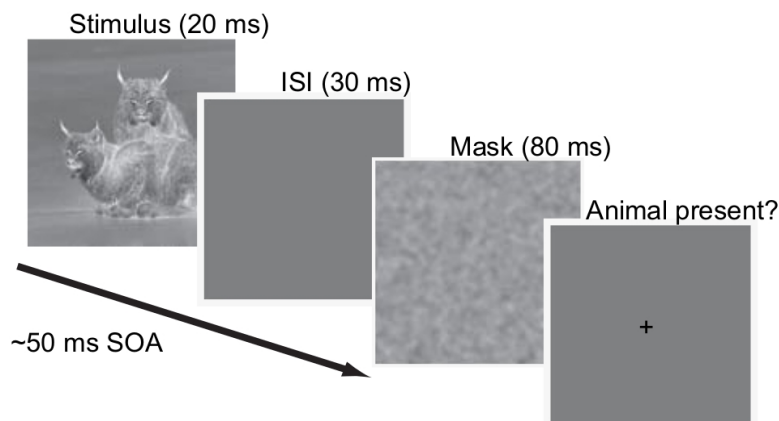


تسک تشخیص تصاویر حیوان / غیر حیوان

تمرین اول و دوم درس علوم اعصاب

علی اکرامیان - 99100563

در این آزمایش ما قصد داریم مطابق آنچه در مقاله‌ی گفته شده¹ انجام داده شده بود را، با چند سوژه (Subject) خودمان تست کرده و نتایج مقاله را باز تولید کنیم. نحوه‌ی کار مانند مقاله می‌باشد:



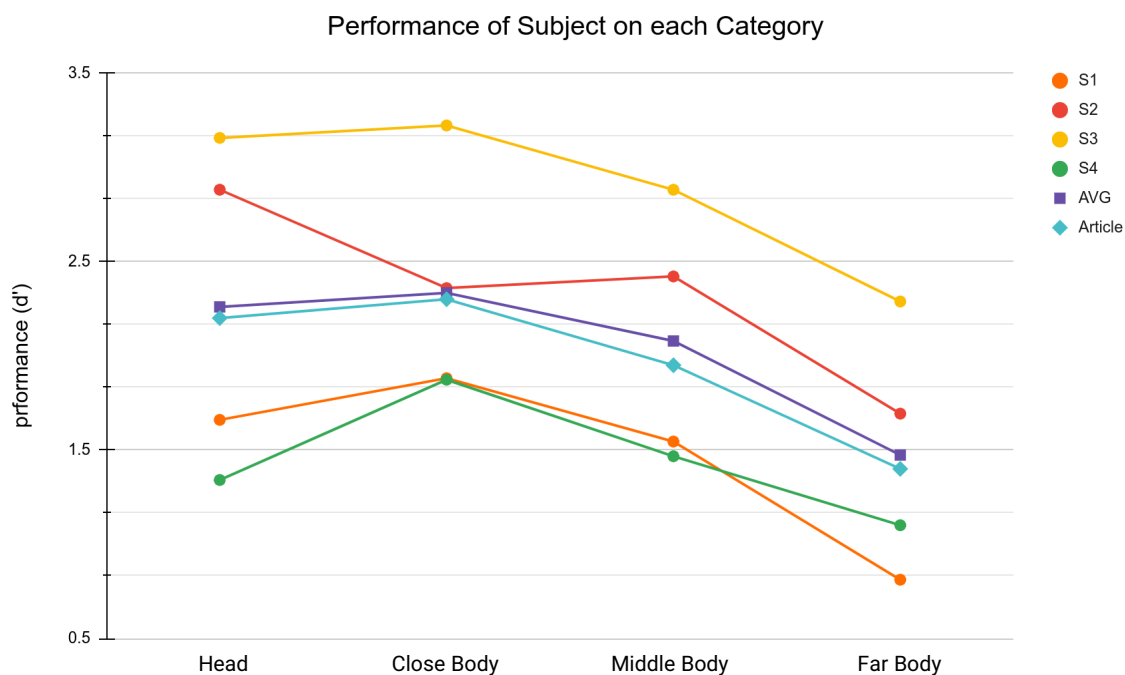
حال کد پایتونی با کتابخانه‌ی psychopy زدم که بتوانم این کار را انجام دهم و نتایج را در فایل CSV ای ذخیره کنم و تحلیل‌هایش را انجام داده و در این گزارش بیاورم. ابتدا من عکس‌ها را از دیتابیس داخل مقاله‌ی کامل (با اضافات) برداشتم و سیاه و سفیدشان کردم. سپس در یک فولدر گذاشتم همه را که کل فولدر را تحویل می‌دهم. حال طبق دستورالعمل مقاله و با در نظر گرفتن تایم‌ها، چند تابع نوشتم که محرک (Stimulus) را لود کرده و سپس تصویر ISI خاکستری را نمایش داده و پس از آن نویزی نمایش می‌دهم (نویز نمایش داده شده نیز در فولدر موجود است). حال از سوژه‌ها (که 4 تا هستیم!) می‌خواستم تا این کار را با فشردن دکمه‌های y یعنی حیوان است و n یعنی حیوان نیست انجام دهند. نتایج را نیز به صورت 1 (درست بودن تشخیص) و 0 (اشتباه بودن تشخیص) در فایل CSV ذخیره می‌کنم. در فایل اکسل که در فولدر موجود است نیز همه‌ی نتایج 4 سوژه را آورده‌ام. در این گزارش نیز نتایج و نمودارها را می‌آورم و دیتاهای کامل در فایل اکسل موجود است. همچنین مدت زمان واکنش را نیز من در کنار درستی / نادرستی ثبت می‌کنم و تحلیل‌های این را نیز می‌آورم. همچنین در فایل، عکس‌ها را سورت کردم تا بتوانم عملکرد (performance) سوژه‌ها را در عکس‌های ثابت مقایسه کنم. از مقایسه می‌توان دید که چه عکس‌هایی است که مثلاً همه جواب نداده‌اند یا همه جواب داده‌اند. مثلاً عکس‌های رو به رو را هیچ‌کدام از سوژه‌ها نتوانستند تشخیص

دهند که Animal است که عکس‌های سختی برای تشخیص هستند! نکته اینجاست که این اتفاق که هیچ‌کس نتواند تشخیص دهد غالباً در نمای دور (Far Body) اتفاق افتاده است که قابل پیش‌بینی نیز هست.

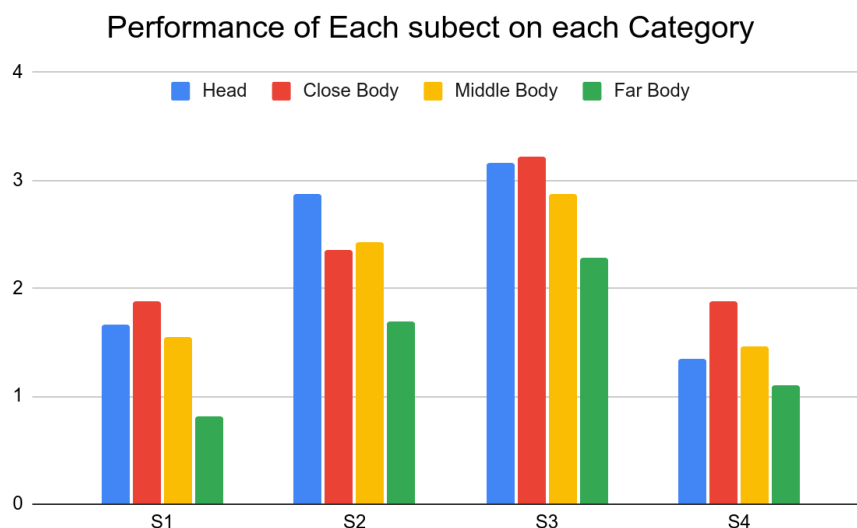


1 Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences of the United States of America*, 104(15), 6424–6429. <https://doi.org/10.1073/pnas.0700622104>

حال کمیت‌های $Z(H)$ و $Z(F)$ را محاسبه می‌کنیم و اختلاف آن‌ها نیز که همان عملکرد (performance) است را بدست آورده و بیشتر با همین کمیت کار می‌کنیم ($d' = Z(H) - Z(F)$). داده‌ها به طور کامل در فایل اکسل Image Rec Task در شیت Per آمده است (شیت اول). اینجا من فقط نتایج را می‌آورم.



نمودار بالا، عملکرد هر سوژه در هر کتگوری (سر، بدن نزدیک، بدن میانه، بدن دور) نشان می‌دهد. رنگ‌های نارنجی، قرمز، زرد و سبز مربوط به سوژه‌ها هستند و نمودار بنفش رنگ میانگین آن‌هاست. همچنین آبی کم‌رنگ نیز عملکرد گزارش شده در عکس مقاله است که اعدادش را به صورت حدودی از عکس صفحه‌ی 4 ام مقاله خواندم. می‌بینیم که با اینکه در هر سوژه روند و میانگین عملکرد هر سوژه به مقاله نزدیک نیست ولی به طور میانگین با تقریب خوبی به عملکرد مقاله نزدیک است که چیز جالبی است! روند نیز مورد انتظار است، ما در تشخیص بدن نزدیک از همه بهتر می‌کنیم و در بدن دور از همه بدتر تشخیص می‌دهیم! البته این میانگین است و در هر سوژه روند حفظ نمی‌شود. مثلاً سوژه دوم، سر را بهتر تشخیص می‌دهد! ولی مثلاً بد تشخیص دادن در بدن‌های دور در همه‌ی سوژه‌ها مشترک می‌باشد. حال به طور مقایسه‌ای نیز می‌توان این‌ها را کار هم دید:

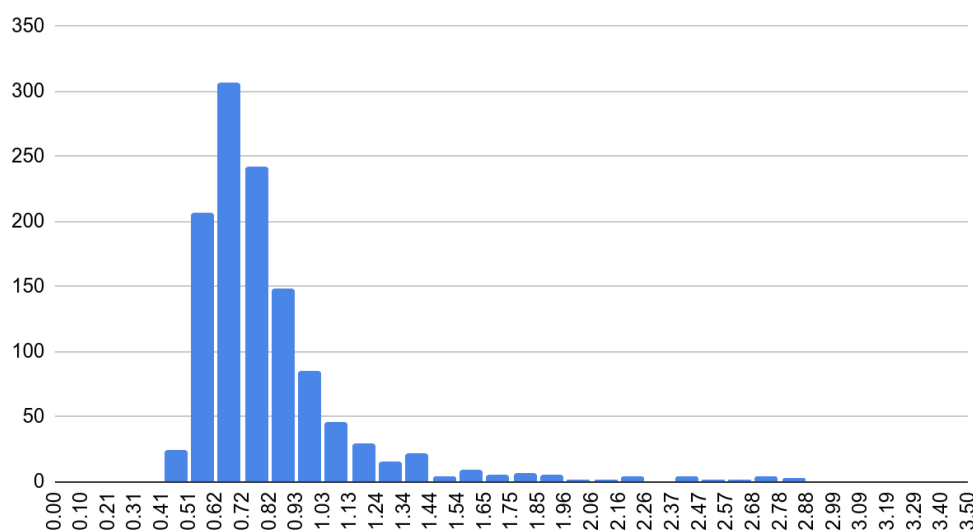


حال زمان واکنش را در شیت دوم در همین فایل اکسل بررسی کردم. در ستون آخر نیز میانگین همه‌ی زمان‌های واکنش به ازای هر عکس را گرفته و در آن ستون آورده‌ام. بدیهی است که ممکن است یک سوژه زمان زیادی را به طور موضعی صرف تشخیص یک عکس کند که می‌توان داده‌های پرت در نظر گرفت و از میانگین کنار زد. ولی فعلا این کار را نمی‌کنیم و کل را در نظر می‌گیریم و نمودار هیستوگرام را با میانگین رسم می‌کنم. می‌توان مانند قبلا بررسی کرد که چه عکسی را مثلا برای همه طول کشیده تا بزنند و تشخیص دهند. مثلا عکس رو به رو یکی از عکس‌هایی است که همه وقت زیادی صرف تشخیص آن کرده اند ولی جالب است که همه نیز آن را درست تشخیص داده‌اند!

حال داده‌های مربوط به زمان واکنش را می‌بینیم.



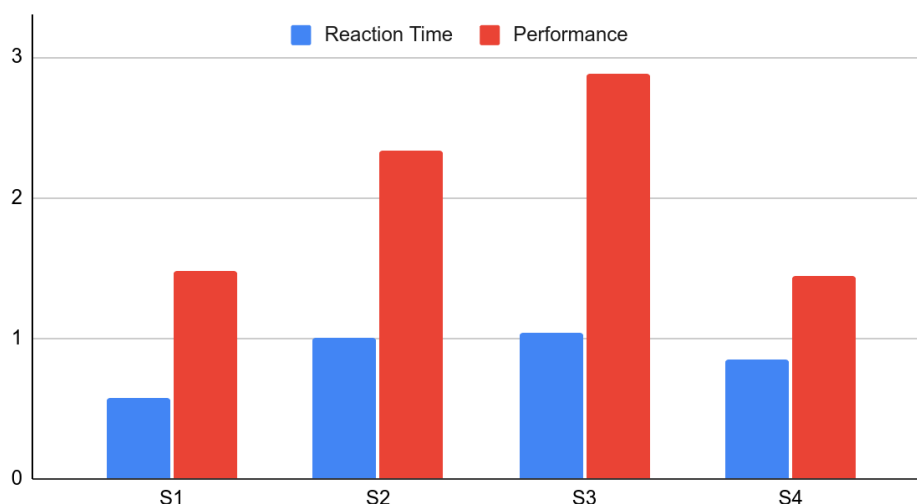
Hisogram of Average Reaction Time of Subjects



این نمودار هیستوگرام مربوط به زمان واکنش است. با تقریب خوبی می‌توان دید که این یک تابع توزیع پواسونی می‌باشد. که البته برای متغیرهایی که توزیع‌های نسبتا رندوم و با فاصله‌ی زمانی هستند، قابل پیش‌بینی می‌باشد.

حال سوالی که می‌توانیم بپرسیم این است که آیا زمان واکنش با عملکرد سوژه کورلیشن دارد یا خیر؟ من این سوال را با رسم نمودار میانگین عملکرد به همراه میانگین زمان واکنش نشان می‌دهم:

Compare of Reacton Time and Performance in Subjects



که نوارهای قرمز رنگ مربوط به عملکرد و نوارهای آبی رنگ زمان واکنش می‌باشد. با اینکه مشاهده می‌کنیم که گویی مقداری تناسب وجود دارد و با مقایسه‌ی سوژه‌های 1 و 2 می‌توان این را دید ولی با مقایسه‌ی 1 و 4 و همچنین 2 و 3 که زمان واکنش تقریباً یکسان ولی عملکرد متفاوتی داشتند، این تناسب را می‌توان رد کرد. البته هیچ کدام قوی نیست و باید تست سوژه‌ها خیلی زیادتر شوند تا بتوان نتیجه‌گیری از داده‌ها انجام داد و 4 داده بسیار کم است برای چنین نتیجه‌گیری‌هایی!

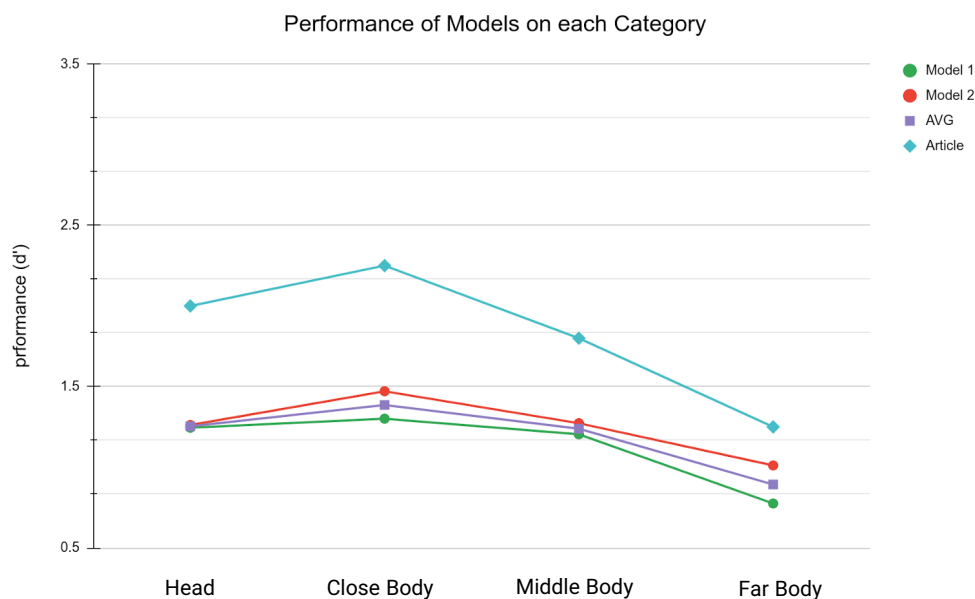
همچنین می‌توان زمان واکنش را بر اساس میزان کتگوری‌های مختلف عکس‌ها نیز دید:



می‌توان دید که در بدن‌های دور خیلی سوژه زمان واکنش زیادی دارد! در بقیه‌ی موارد خیلی اعداد اختلاف کمی دارد. از اردر 10 میلی‌ثانیه است این اختلاف کم است و احتمالاً خیلی قابل اعتماد نیست نتیجه‌گیری در این شرایط.

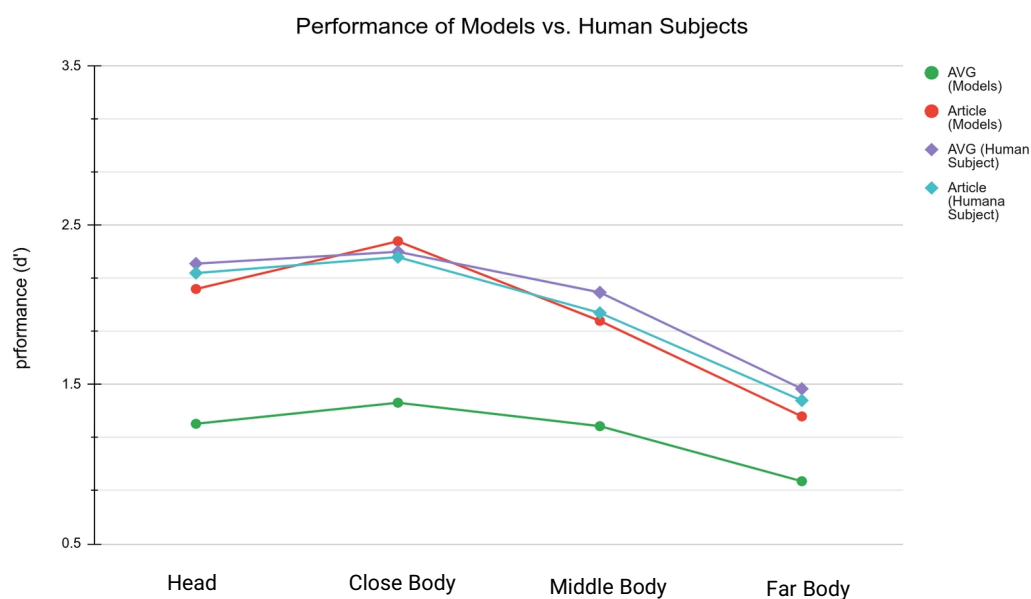
مدل کامپیوتری

حال با مدل کامپیوتری در کد متلب انجام می‌دهیم و با نتایج سوژه‌های خودمان مقایسه می‌کنیم. در کد متلب من از مدل SVM استفاده نکردم. حال همان عکس‌هایی که در مرحله‌ی قبل سیاه و سفید کردم را نیز به این کد می‌دهم. این عکس‌ها را نصف کردم و در دو مدل نصفه‌ی اول را نمونه‌گیری کرده و برای ترین مدل اول استفاده کردم که این عکس‌ها همان مدل‌های تست مدل دوم هستند. حال از نیمه‌ی دوم نیز همین استفاده را کردم و بالعکس این کار را انجام دادم. قابل ذکر است که این نیمه‌ها در کتگوری‌های مختلف بودند (مثلاً نصف کتگوری سر را به مدل 1 برای تست دادم و نصف دیگر برای ترین آن و بالعکس برای مدل دوم). حال همان روش ذخیره‌سازی درست و غلط را که برای مدل انسانی (سوژه‌ها!) استفاده کرده بودم را اینجا نیز برای مدل‌ها استفاده می‌کنم. نتایج در شیت سوم در فایل اکسل موجود است که برای داده‌های تست هر مدل، چه میزان 1 (تشخیص درست) و 0 (تشخیص غلط) داریم. حال همان نمودار عملکرد مدل را به همراه میانگین و مقایسه با نتایج مقاله در یک شکل رسم می‌کنم.



قابل مشاهده است که عملکرد مدل به دلایلی مثل کم بودن تست‌ها و استفاده نکردن از مدل SVM و غیره با نتایج مقاله تفاوت دارد و مقداری عملکرد ضعیف‌تر می‌باشد. چیزی که مورد توجه است روندی است که کتگوری‌های مختلف عکس‌ها دارند. می‌توان در این مدل‌ها نیز دید که عملکرد مدل، مانند سوژه‌ها در بدن نزدیک، بیشترین مقدار است. سپس در سر عملکرد خوب است و از همه بدتر بدن دور است که از همه ضعیف‌تر است (مانند انسان!).

برای بهتر کردن نتایج مدل‌ها نیز باید تعداد عکس بیشتری به آن‌ها داد تا بتواند بهتر ترین شود. البته با توجه به زمان زیادی که می‌برد دیگر من بیشتر ندادم و به همین میزان عکس کفایت کردم. حال که میانگین نتایج دو مدل کامپیوتری و چهار سوژه‌ی انسانی را داریم می‌توانیم آن‌ها را مقایسه نیز بکنیم. دو میانگین را با هم و با نتایج مقاله یک‌جا رسم می‌کنیم:



که همان‌طور که می‌بینیم، طبق مقاله، عملکرد مدل فقط در بدن نزدیک بهتر از انسان بوده و در بقیه‌ی موارد ضعیف‌تر عمل کرده است. البته مدل ما به دلیل کم‌بودن ترینینگ خیلی نمی‌توان با سوژه‌های انسانی مقایسه‌ای داشت چون تفاوت دارند.

حال می‌توانیم ببینیم چه عکس‌هایی هستند که هم مدل و هم سوژه‌های انسانی آن‌ها را نتوانستند تشخیص دهند! مثلا عکس‌های زیر عکس‌هایی هستند که هیچ‌یک از سوژه‌های انسانی (چهار نفر) و همچنین مدل کامپیوتری نیز نتوانستند تشخیص درستی داشته باشید. این موارد که هیچ‌یک از سوژه‌های انسانی و همچنین مدل نتوانستند تشخیص دهند اکثرشان با درصد بالا در کتگوری بدن دور (Far Body) بودند.



همچنین عکس‌های پایین نیز عکس‌هایی است که مثلا سوژه‌های انسانی هیچ‌کدام نتوانستند درست پاسخ دهند ولی مدل توانسته بود درست تشخیص دهد.

