Assignment 7

Neural Networks and Deep Learning

CSCI 5922

Fall 2017

Art Image Similarity Judgement using Autoencoder with Dropout

Student: **Bo Cao**

bo.cao-1@colorado.edu
boca7588@colorado.edu

Github:
https://github.com/BryanBo-Cao/neuralnets-deeplearning
https://github.com/tylersco

Introduction

Images in Cubist and Impressionist look similar from our human's eyes, but the features among them are different. In this paper our raw data from participants show that human tend to judge two images as in the same category if they are similar. Autoencoder is used to be trained to learn the hidden representation of an image, and then the correlation of image's similarity and the probability of human's judgement on the same category is evaluated based on the image representation after the autoencoder's bottleneck and after the decoder. Our results show that some settings of the autoencoder outperform baseline, which shows a slightly stronger correlation in many of the Beta set. In the other group of the experiment, 25% dropout is introduced in the bottleneck layer and slightly boost the performance.

The Dataset

The dataset contains 2592 images with colors, and a table of metadata in the csv format including the result from human judgement on image pairs. All 2592 images are in the shape of 16 * 16 with RGB colors, which are used for training the model. The values from the original image are from 0 to 255. When loading these images into the network, they are normalized to [0, 1].

These images are sampled from 512x512 pixel images, which are shown to human to judge image pair similarity. In terms of the human judgement experiment, participants are given pairs of images and judge whether they are in the same category or not. The category includes Cubist and Impressionist. The metadata table contains the data from these experiments, including 96 image pairs in the same category of Cubist, 96 image pairs in the same category of Impressionist and 192 image pairs in different categories of Cubist and Impressionist.

In the metadata table, each image pair contains both cosine value and accuracy of human's judgement. The similarity of an image pair is measure by their cosine value, which is calculated by the two images(image A and B) flatten one dimensional vector's function:

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

Figure 1

The greater the cosine value is, the more similar the two images are. Likewise, in my code, distance of these two images are stored as well, using Euclidean distance function. The greater the distance is, the more different the two images are.

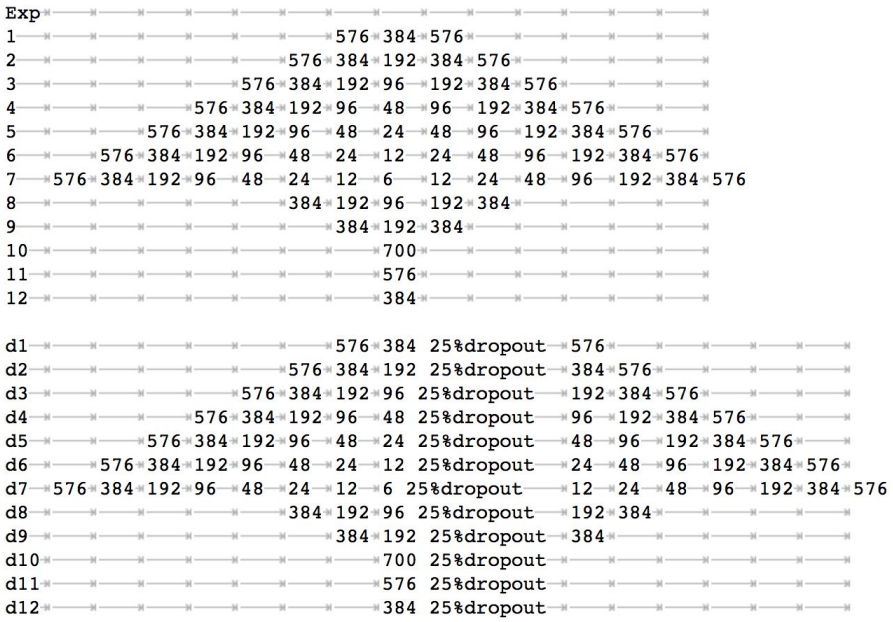
In addition to human's judgement accuracy, a new feature of probability of human's judgement as similar is added, which denotes the probability in the case that human will judge two images in the same category. To be more specific, if two images are in the same category, then this variable is identical to human's judgement accuracy; on the contrary, if two images are in different categories, then this variable equals (1 - human's judgement accuracy).

All of the mage pairs are equally splitted into Alpha and Beta set. In particular, each set contains 48 image pairs in the same category of Cubist, 48 image pairs in the same category of Impressionist and 96 image pairs in different categories. Alpha set is used for model selection while Beta set is used to evaluate the network.

The Architecture

Autoencoder

Autoencoder is used to construct the Architecture. A total number of 24 experiments were done to measure the performance of the neural network. All of these experiments are divided into 2 groups - pure autoencoder and autoencoder with dropout (25%) in the bottleneck layer.



The figure depict each experiment's setting and the core architecture.

Group 1 contains experiments 1 to 12, no dropout is used. All the number in the graph denotes the number of neuron in the corresponding hidden layer. Each layer is fully connected with its next layer and sigmoid is used as the activation function. The number

in the middle in each row represents the bottleneck. In each row, numbers from left to the middle represents encoder, while numbers from middle to the right represents decoder. For instance, in experiment 2, the autoencoder consists of 5 hidden layers, the bottleneck contains 192 neurons.

Group 2 contains experiments d1 to d12, which are almost identical to group 1, except there's 25% dropout in the bottleneck.

Each layer is one dimensional vector. To feed 2D images, each image is flattened from 16x16x3 to one dimensional vector with length 768.

Results

Although in the first betting, alpha set is used only to select model and beta set is used for evaluation, all of the results were collected, including alpha, beta and alpha & beta.
Each neural networks in each experiment setting is trained with 1000 epochs, batch size of 256 and learning rate as 0.01.

The correlation between the [distance, cosine value] and [probability of human judging the same category, accuracy] is calculated using Spearman correlation equation, after which we can get rho and p-value. Rho denotes how strong the correlation between two variables is, the larger it is, the stronger the correlation is. The sign denotes if the two variables is positive or negative relative. P-value illustrates how strong we can believe the hypothesis, in this project we set alpha value as 0.05, which means that any p-value lower than 0.05 denotes significance.

In each experiment, two outputs are used to measure their performance, including the one after bottleneck and the one after the whole network, so that we could see whether the hidden representation can strengthen the correlation.

The following figures(3-7) show network each setting's performance compared to baseline result. For those whose result is better than the baseline, the cell is displayed in green. In particular, for those rho value whose absolute value is greater than the rho value in baseline, or the p-value is smaller than the one in baseline, then it is considered as outperformance. If the p-value is smaller than 0.05, then it is considered as significant, denoted in bold.

		Alpha							
		Distance & Probability Same		Cosine & Probability Same		Distance & Accuracy		Cosine & Accuracy	
Exp	Output	rho	p-value	rho	p-value	rho	p-value	rho	p-value
baseline		**-0.5068243769**	**6.34E-14**	**0.4857974121**	**9.17E-13**	**-0.09097439599**	**0.2094966761**	**0.07441810071**	**0.3049589243**
1	bottlenect	-0.178637212	**0.01317217747**	-0.1804668233	**0.01224881669**	-0.01494332219	0.8370095367	-0.01546741124	0.8313769231
1	network	-0.1812553027	**0.01186876174**	-0.1977675872	**0.00596562142**	-0.02594155984	0.720964793	-0.01882141154	0.795544908
2	bottlenect	-0.316267561	**7.86E-06**	-0.3133934264	**9.59E-06**	0.01963637849	0.7868994909	0.01978478559	0.7853279104
2	network	-0.3402025735	**1.38E-06**	-0.343803296	**1.05E-06**	0.03240108456	0.6554903094	0.04676434703	0.5195060707
3	bottlenect	-0.2266734111	**0.001568916319**	-0.2280172174	**0.00146830446**	0.00899431787	0.9014596961	0.008884920642	0.9026522803
3	network	-0.2430653041	**0.000681016073**	-0.2556852135	**0.000344348630**	-0.00539438581	0.9408041053	-0.01012645198	0.8891312733
4	bottlenect	-0.2347472705	**0.001047556212**	0.01948646265	0.7884879234	0.01948712336	0.7884809209	-0.08092850786	0.2644704692
4	network	-0.200727148	**0.005243762679**	-0.1902427905	**0.008216281116**	0.09893082054	0.1721821319	0.1007695754	0.1643076462
5	bottlenect	-0.3018265186	**2.09E-05**	-0.03805133702	0.6002773938	0.005746322637	0.9369498278	-0.05614537139	0.4392245669
5	network								
6	bottlenect	-0.2834227314	**6.79E-05**	0.02333305511	0.7480248767	-0.1197518034	0.09803545558	-0.04659728305	0.5210008336
6	network	-0.09792372276	0.176612041	-0.09792372276	0.176612041	0.08358259278	0.2490706369	0.08358259278	0.2490706369
7	bottlenect	-0.1859412442	**0.009816934855**	0.01785948206	0.8057814072	-0.00312757347	0.9656584531	0.09601006066	0.1852608535
7	network	-0.01760487321	0.8084964739	0.04357351669	0.5484296975	-0.01122221305	0.8772241864	-0.04202788717	0.5627183732
8	bottlenect	-0.4217194726	**1.12E-09**	-0.4216177334	**1.13E-09**	0.0509858928	0.4824765685	0.05104101543	0.4820026557
8	network	-0.3938115414	**1.59E-08**	-0.4279077641	**5.98E-10**	0.02549973072	0.7255249864	0.04362574903	0.5479499551
9	bottlenect	-0.3646768041	**1.99E-07**	-0.3644495864	**2.03E-07**	0.01653170211	0.8199652199	0.01651558934	0.8201377115
9	network	-0.368059635	**1.50E-07**	-0.3728888593	**1.00E-07**	-0.00509163534	0.944120817	-0.00216928767	0.9761768853
10	bottlenect	-0.4614308555	**1.63E-11**	-0.4607135936	**1.77E-11**	-0.08235151987	0.2561345919	-0.08296126101	0.252618797
10	network	-0.5049218525	**8.13E-14**	-0.5151881933	**2.08E-14**	-0.03554307473	0.6245316646	-0.02509691147	0.729690988
11	bottlenect	-0.4700625852	**6.03E-12**	-0.4701524549	**5.97E-12**	-0.01928020147	0.7906747879	-0.01991368775	0.7839635843
11	network	-0.4522522772	**4.54E-11**	-0.4795421423	**1.96E-12**	-0.01208203344	0.8679000713	-0.00082344735	0.9909557609
12	bottlenect	-0.4603795497	**1.83E-11**	-0.4599463099	**1.93E-11**	-0.01848134729	0.7991598562	-0.01937942221	0.7896226042
12	network	-0.4320358352	**3.92E-10**	-0.4790436198	**2.08E-12**	-0.00937254394	0.8973382143	-0.01364327606	0.8510169387

Figure 2

		Alpha							
		Distance & Probability Same		Cosine & Probability Same		Distance & Accuracy		Cosine & Accuracy	
Exp	Output	rho	p-value	rho	p-value	rho	p-value	rho	p-value
baseline		**-0.5068243769**	**6.34E-14**	**0.4857974121**	**9.17E-13**	**-0.09097439599**	**0.2094966761**	**0.07441810071**	**0.3049589243**
d1	bottlenect	-0.3261879882	**3.89E-06**	-0.3269086414	**3.69E-06**	0.03265040848	0.6530111869	0.03214582436	0.658032366
d1	network	-0.2722933024	**0.000133009529**	-0.2995331457	**2.44E-05**	-0.04987411166	0.4920884429	-0.03832803983	0.5976278204
d2	bottlenect	-0.2732394777	**0.000125756972**	-0.1018249065	0.1599114551	-0.00865001341	0.9052138339	0.006984461812	0.9234024396
d2	network	-0.08084146441	0.264986315	-0.00746190912	0.9181840067	-0.1103487261	0.1275813499	-0.08011908504	0.2692938922
d3	bottlenect	-0.1418508333	**0.04968759847**	-0.07046631816	0.3314278195	0.1119532226	0.1221036688	0.05805939887	0.4237592741
d3	network	-0.08852354988	0.2220833421	-0.1092413458	0.1314711728	0.09236935192	0.2025657895	0.1336699317	**0.06454539385**
d4	bottlenect	-0.279914422	**8.42E-05**	-0.1575169861	**0.02910902586**	-0.06092407978	0.4012132499	0.1289996854	**0.07454427941**
d4	network	-0.3060470969	**1.58E-05**	-0.277708143	**9.62E-05**	-0.007538012611	0.9173525233	-0.04015575461	0.5802605403
d5	bottlenect	-0.1509174984	**0.03666258147**	-0.09049878159	0.2118982889	-0.0322993197	0.6565032814	0.004492070692	0.9506919889
d5	network	-0.02057472827	0.7769774324	0.02728026878	0.7072083188	-0.01231246681	0.8654042316	0.0245660694	0.7351931091
d6	bottlenect	-0.4163179649	**1.90E-09**	-0.03925016482	0.588836304	0.03898781537	0.5913315036	-0.0741229826	0.3068861379
d6	network								
d7	bottlenect	-0.09637693166	0.1835791425	-0.11795906	0.1032052046	-0.01996502805	0.7834203714	0.03168788248	0.6626026957
d7	network								
d8	bottlenect	-0.4260060875	**7.25E-10**	-0.42382378	**9.04E-10**	0.02329058514	0.7484681137	0.02342881575	0.7470257758
d8	network	-0.4036717729	**6.40E-09**	-0.4254202388	**7.70E-10**	0.05104610368	0.4819589225	0.06422126136	0.3761679625
d9	bottlenect	-0.4070766474	**4.64E-09**	-0.4067019077	**4.81E-09**	0.07238110735	0.3184238168	0.07160854243	0.3236301327
d9	network	-0.4119745457	**2.90E-09**	-0.4493433817	**6.25E-11**	0.07676462885	0.2899193775	0.08525521065	0.2396913094
d10	bottlenect	-0.4735301989	**4.02E-12**	-0.4728392197	**4.36E-12**	-0.0120498079	0.8682492131	-0.01327861863	0.8549545446
d10	network	-0.4792157288	**2.04E-12**	-0.4925834218	**3.95E-13**	0.02164199435	0.7657360169	-0.00028154945	0.996907576
d11	bottlenect	-0.4877194366	**7.24E-13**	-0.4875100232	**7.43E-13**	-0.05133782962	0.4794551397	-0.05103253503	0.482075549
d11	network	-0.4694038234	**6.51E-12**	-0.4936220102	**3.47E-13**	-0.02875111812	0.6922027283	-0.02270458913	0.7545922116
d12	bottlenect	-0.4141136139	**2.36E-09**	-0.413774483	**2.44E-09**	-0.03467468123	0.6330255881	-0.03538957939	0.6260294781
d12	network	-0.4486379894	**6.75E-11**	-0.4617250516	**1.57E-11**	-0.01986958964	0.7844302535	-0.01186069486	0.8702986357

Figure 3

		Beta							
		Distance & Probability Same		Cosine & Probability Same		Distance & Accuracy		Cosine & Accuracy	
Exp	Output	rho	p-value	rho	p-value	rho	p-value	rho	p-value
baseline		**-0.2080056978**	**0.003790081657**	**0.09716503023**	**0.1800045387**	**-0.02214295401**	**0.7604761261**	**-0.09742611325**	**0.1788317251**
1	bottlenect	-0.09607118275	0.1849798921	-0.09646293277	0.1831865476	0.000535312975	0.9941203795	3.48E-05	0.999617961
1	network	-0.09062483982	0.2112598445	-0.07611821757	0.2940120798	-0.008819514575	0.9033654031	0.01686193456	0.8164319058
2	bottlenect	-0.1237565473	0.08722818845	-0.123142636	0.08882010129	0.04331199472	0.5508348091	0.04310499573	0.5527421379
2	network	-0.08246168502	0.2554968916	-0.06368990554	0.380138108	0.08275378691	0.2538113377	0.1098562333	0.129300188
3	bottlenect	-0.1086818027	0.1334710154	-0.1109356373	0.1255560736	0.1530511548	**0.03405723184**	0.1521815893	**0.03509943631**
3	network	-0.1288340347	0.07492078046	-0.1317187394	0.06858052767	0.1722910327	**0.0168647672**	0.1826299538	**0.01123092644**
4	bottlenect	0.04934863232	0.4966666755	0.04961617962	0.494332862	0.1682961218	**0.01962610069**	0.166955365	**0.02063679381**
4	network	-0.07721949789	0.287062326	-0.07721949789	0.287062326	0.002695253871	0.9704030977	0.002695253871	0.9704030977
5	bottlenect	-0.03566621143	0.6233312031	-0.03777966572	0.602883903	0.1500963289	**0.03770922707**	0.1537263845	**0.03326615553**
5	network								
6	bottlenect	-0.003493528385	0.9616430872	-0.001092999537	0.9879953599	-0.04694126399	0.5179255479	-0.0432051018	0.55181934
6	network								
7	bottlenect	-0.1949185334	**0.006743149894**	-0.191229978	**0.00788339487**	0.0326252474	0.6532612013	0.03561994599	0.6237821294
7	network	-0.09385011042	0.1953915876	0.00826276175	0.9094387568	0.02328002509	0.7485783363	-0.01210493346	0.8676519808
8	bottlenect	-0.2422439719	**0.000711067510**	-0.2414138349	**0.000742679098**	-0.07095484418	0.3280781403	-0.06986894781	0.3355535162
8	network	-0.242559407	**0.000699384463**	-0.185437674	**0.01002127134**	-0.1254710837	0.08290301602	-0.104535341	0.1490255368
9	bottlenect	-0.1438324643	**0.04655260474**	-0.1433966212	**0.04722756643**	0.09601444797	0.1852406758	0.09526025901	0.1887330032
9	network	-0.2217042296	**0.001998086535**	-0.2158118683	**0.002643973488**	0.04052684021	0.5767632814	0.0473510202	0.5142744082
10	bottlenect	-0.2213497891	**0.002032448281**	-0.2219832031	**0.001971413975**	0.01655652621	0.8196994881	0.01558091619	0.8301581567
10	network	-0.2702752975	**0.00014980336**	-0.228169801	**0.001457261382**	-0.02785748366	0.7013057315	0.02825196945	0.6972819658
11	bottlenect	-0.2264849367	**0.001583520433**	-0.2264111656	**0.001589270385**	-0.0638473131	0.3789593473	-0.06546852244	0.3669490037
11	network	-0.2469754982	**0.00055335931**	-0.1968832192	**0.006197926002**	-0.07949864288	0.2730314744	-0.01896585878	0.7940107179
12	bottlenect	-0.2680791231	**0.000170318768**	-0.2683131556	**0.000168013880**	-0.03657858887	0.6144675007	-0.0386842098	0.5942250916
12	network	-0.2880465033	**5.09E-05**	-0.204518953	**0.004433721046**	-0.06378368635	0.3794355529	0.001186002441	0.9869739762

Figure 4

		Beta							
		Distance & Probability Same		Cosine & Probability Same		Distance & Accuracy		Cosine & Accuracy	
Exp	Output	rho	p-value	rho	p-value	rho	p-value	rho	p-value
baseline		**-0.2080056978**	**0.003790081657**	**0.09716503023**	**0.1800045387**	**-0.02214295401**	**0.7604761261**	**-0.09742611325**	**0.1788317251**
d1	bottlenect	-0.05598040065	0.4405724104	-0.05643489852	0.4368647618	0.1268157288	0.07963300629	0.1274155169	0.07820833827
d1	network	-0.1624558536	**0.02436453601**	-0.1316975408	0.06862547133	0.07380108036	0.3089973664	0.1006099953	0.1649801948
d2	bottlenect	0.07017837137	0.3334124252	0.07094663851	0.3281342236	0.09194148823	0.2046737894	0.09202593871	0.2042564698
d2	network	-0.0180129963	0.8041455004	0.000188322627	0.9979315385	0.0248954287	0.7317777276	0.09587596101	0.1858783697
d3	bottlenect	-0.05358495947	0.460406873	-0.05431301129	0.4543266906	-0.1404123391	0.05207175714	-0.1402602969	0.05232920787
d3	network	-0.1320518077	0.0678775271	-0.05106446444	0.4818011307	-0.2636329747	**0.000220120151**	-0.2047312155	**0.004391894078**
d4	bottlenect	-0.03665237005	0.6137531324	-0.03658029483	0.6144509792	0.1114774426	0.123708592	0.1115427661	0.1234872775
d4	network	0.005164297556	0.9433246934	0.001611151314	0.9823051685	0.1210629492	0.09438610744	0.1144465567	0.1139556476
d5	bottlenect	-0.1828378784	**0.01113712674**	-0.1827810661	**0.0111626873**	0.1040661999	0.1508684713	0.102034386	0.159049421
d5	network	-0.04219314399	0.5611821721	0.03404068995	0.6392573211	0.05688817814	0.4331848718	-0.01796064511	0.8047032796
d6	bottlenect	-0.04217163614	0.5613819917	-0.04127495308	0.5697428985	0.2373922168	**0.000914968802**	0.234793376	**0.001045101335**
d6	network								
d7	bottlenect	0.1442078095	**0.04597780561**	0.1414100029	0.05040838196	0.1342321237	0.0634195267	0.1357886495	0.06038611878
d7	network								
d8	bottlenect	-0.2594165294	**0.000279589702**	-0.2603950565	**0.000264588738**	-0.06161868047	0.3958561197	-0.06245091823	0.3894942221
d8	network	-0.2435633724	**0.000663367486**	-0.1961132862	**0.006406715171**	-0.03922546128	0.5890710531	0.0173862189	0.8108299615
d9	bottlenect	-0.2079989142	**0.003791247448**	-0.2081718948	**0.003761620561**	-0.0188776297	0.7949477152	-0.01864348329	0.7974357817
d9	network	-0.2619298348	**0.000242560549**	-0.2052897339	**0.004283526264**	-0.08342738201	0.2499537234	-0.04427148597	0.5420360584
d10	bottlenect	-0.2056568935	**0.004213603881**	-0.2071976074	**0.003931252476**	-0.08123862214	0.2626382224	-0.08228210068	0.2565369938
d10	network	-0.2277636359	**0.001486826501**	-0.1528732409	**0.03426830572**	-0.07350500393	0.310947603	-0.0322740278	0.6567551354
d11	bottlenect	-0.2313572217	**0.001243257575**	-0.231035851	**0.001263449167**	-0.08077541947	0.265378176	-0.0818172013	0.259243097
d11	network	-0.2685599072	**0.000165615273**	-0.1915403254	**0.007781248775**	-0.05809969755	0.4234370963	-0.007427361495	0.9185614929
d12	bottlenect	-0.237536188	**0.000908214931**	-0.2389369911	**0.000844853297**	0.006599365514	0.9276138835	0.004388548374	0.951826933
d12	network	-0.2749381404	**1.14E-04**	-0.1921237107	**0.007592416154**	-0.04036989425	0.5782411996	0.02731029369	0.7069008501

Figure 5

		All							
		Distance & Probability Same		Cosine & Probability Same		Distance & Accuracy		Cosine & Accuracy	
Exp	Output	rho	p-value	rho	p-value	rho	p-value	rho	p-value
baseline		**-0.4531196797**	**1.65E-30**	**0.4270180331**	**6.26E-27**	**0.1846075248**	**8.22E-06**	**-0.2314204771**	**1.93E-08**
1	bottlenect	-0.1565670969	**0.002090449466**	-0.1570051743	**0.002030123475**	-0.02444430763	0.632993027	-0.02456557619	0.6313069764
1	network	-0.1552425858	**0.002282940583**	-0.1616789535	**0.001478565949**	-0.02791890635	0.5854651408	-0.02006719095	0.6950637408
2	bottlenect	-0.2581011203	**2.92E-07**	-0.2566781636	**3.41E-07**	-0.02934338794	0.5664690102	-0.02950472601	0.5643361958
2	network	-0.2649448342	**1.37E-07**	-0.2654011383	**1.30E-07**	-0.01828589994	0.7209490655	-0.00246014055	0.9616751032
3	bottlenect	-0.1818756559	**0.000340551567**	-0.1830466175	**0.000311212176**	0.05755804541	0.2605218067	0.05740009947	0.261834384
3	network	-0.2062577247	**4.65E-05**	-0.2204652503	**1.30E-05**	0.05083357643	0.3204538976	0.0471327652	0.3569926095
4	bottlenect	-0.1495556335	**0.003306948935**	-0.01419602863	**0.7815564552**	0.04076065347	0.4257591985	-0.02859659136	0.57639116
4	network	-0.1706559241	**0.000785307336**	-0.1592428907	**0.001746066902**	0.0772348535	0.1308365494	0.07779404482	0.1280634913
5	bottlenect	-0.1958611757	**0.000111981627**	0.04098605497	0.4232025254	0.06067604671	0.2355355191	0.01490793451	0.7708995415
5	network								
6	bottlenect	-0.2102678177	**3.27E-05**	0.01977271903	0.6993195456	-0.1497350852	**0.003269122754**	-0.07925091352	0.1210525306
6	network	-0.07375667028	0.1491379113	-0.07375667028	0.1491379113	0.05902461129	0.2485506328	0.05902461129	0.2485506328
7	bottlenect	-0.1737857422	**0.000625172760**	0.005696171619	0.9114111913	0.0384010047	0.4530580863	0.07558147947	0.1393073537
7	network	-0.02860073373	0.5763358988	0.04380896641	0.3919484027	0.03326209564	0.5157851101	-0.01905905051	0.7096726026
8	bottlenect	-0.3439146933	**4.21E-12**	-0.3435506885	**4.45E-12**	-0.01232359963	0.8097767093	-0.01243765992	0.8080501842
8	network	-0.3373945039	**1.12E-11**	-0.3545769076	**8.09E-13**	-0.0460564007	0.3680921253	-0.0312775367	0.5411575992
9	bottlenect	-0.2969920915	**2.93E-09**	-0.2972987669	**2.82E-09**	-0.01404192428	0.783868794	-0.01346695864	0.7925128001
9	network	-0.3135968562	**3.29E-10**	-0.3148952989	**2.76E-10**	-0.03817606423	0.4557108394	-0.03471779013	0.4975729098
10	bottlenect	-0.4091656516	**6.24E-17**	-0.4092049663	**6.19E-17**	-0.05629278181	0.271164308	-0.05656309198	0.2688660923
10	network	-0.4425567065	**7.57E-20**	-0.4413993091	**9.67E-20**	-0.03662331782	0.474258167	-0.02898445843	0.5712276579
11	bottlenect	-0.3790058148	**1.45E-14**	-0.3789489093	**1.47E-14**	-0.04986491198	0.3297738408	-0.05007755598	0.3277130668
11	network	-0.3808398267	**1.06E-14**	-0.3828522911	**7.48E-15**	-0.03803550294	0.4573729073	-0.03199539571	0.5319085796
12	bottlenect	-0.4040090761	**1.65E-16**	-0.4037824076	**1.72E-16**	-0.02061311149	0.6871992698	-0.02158050389	0.6733460461
12	network	-0.3895409718	**2.31E-15**	-0.3963309593	**6.79E-16**	-0.02820914526	0.5815708493	-0.02498280789	0.6255206449

Figure 6

		All							
		Distance & Probability Same		Cosine & Probability Same		Distance & Accuracy		Cosine & Accuracy	
Exp	Output	rho	p-value	rho	p-value	rho	p-value	rho	p-value
baseline		**-0.4531196797**	**1.65E-30**	**0.4270180331**	**6.26E-27**	**0.1846075248**	**8.22E-06**	**-0.2314204771**	**1.93E-08**
d1	bottlenect	-0.2502670166	**6.78E-07**	-0.2507003253	**6.48E-07**	0.03349832025	0.5128064651	0.03351973306	0.512536905
d1	network	-0.2405799325	**1.85E-06**	-0.2498064737	**7.12E-07**	-0.03447567702	0.5005781577	-0.02263397152	0.6583849086
d2	bottlenect	-0.1635583203	**0.001298497014**	-0.04353603496	0.3949078395	-0.00114971925	0.9820839211	0.007126860499	0.8892890701
d2	network	-0.05092909155	0.3195442854	0.000496187385	0.9922673836	-0.04762558944	0.3519814246	0.02103199692	0.6811875762
d3	bottlenect	-0.1071589835	0.03580981127	-0.01824014086	0.7216183919	0.02555607745	0.6176077074	0.005804567605	0.9097324953
d3	network	-0.08370992719	0.1014423592	-0.0726222457	0.155509846	0.002486408761	0.9612662069	0.04021217587	0.4320178195
d4	bottlenect	-0.186726086	**0.000233643111**	-0.1054900719	**0.03880923148**	-0.02912332366	0.5693843519	0.1075796313	**0.035085664**
d4	network	-0.190507284	**0.000173034365**	-0.1748760835	**0.000576902467**	0.03860946126	0.450607502	0.0266972719	0.6019871818
d5	bottlenect	-0.1574905141	**0.001965150074**	-0.1077165944	**0.03485258479**	0.002482189383	0.9613318859	0.01313389411	0.7975317946
d5	network	-0.03010619734	0.5564191366	0.03269264632	0.5230020947	-0.00371628542	0.942134882	0.03300685629	0.5190135402
d6	bottlenect	-0.2739096311	**4.91E-08**	-0.0143760706	0.7788573577	0.07661862249	0.1339459541	-0.00281196793	0.9561993619
d6	network								
d7	bottlenect	-0.02270316945	0.6574068436	-0.06522955487	0.2021590223	0.03399400777	0.5065853028	-0.01873620609	0.7143737818
d7	network								
d8	bottlenect	-0.3918976247	**1.51E-15**	-0.3914580637	**1.64E-15**	-0.01385005706	0.786750422	-0.01276712032	0.8030684128
d8	network	-0.362256402	**2.37E-13**	-0.3588771447	**4.09E-13**	0.01109607345	0.8284144416	0.0177770808	0.7284035264
d9	bottlenect	-0.3578889803	**4.79E-13**	-0.3574474059	**5.14E-13**	-0.00060994694	0.990494623	-0.00049058645	0.9923546662
d9	network	-0.377431004	**1.90E-14**	-0.3957350935	**7.57E-16**	-0.01280231364	0.8025367215	-0.00321467694	0.9499342549
d10	bottlenect	-0.3942363687	**9.93E-16**	-0.3942658282	**9.88E-16**	-0.05248511856	0.3049631032	-0.05334587094	0.2970891427
d10	network	-0.4096185647	**5.72E-17**	-0.3971918545	**5.81E-16**	-0.04734498518	0.35482923	-0.04874020967	0.3408122177
d11	bottlenect	-0.4212139453	**6.02E-18**	-0.4207677083	**6.58E-18**	-0.07695020297	0.1322658498	-0.07733732953	0.1303249233
d11	network	-0.4159261796	**1.70E-17**	-0.4095421608	**5.81E-17**	-0.04946803305	0.3336423809	-0.04782539525	0.349962441
d12	bottlenect	-0.3653340704	**1.44E-13**	-0.3653029154	**1.44E-13**	-0.00818753602	0.8729425908	-0.00880977943	0.8633775875
d12	network	-0.3972273542	**5.77E-16**	-0.375694166	**2.56E-14**	-0.02975086727	0.56108976	-0.02369634349	0.6434340558

Figure 7

In terms of the baseline, Figure 8 and 9 show the graph, the y-axis is human's judgement accuracy. In Figure 8, the x-axis is two images' distance while in Figure 9, the x-axis is two images' cosine value. In this paper image pairs in alpha set are denoted in red while those in beta set are denoted in blue. In Figure 8, rho = 0.18, p-value = 8.22e-06. In Figure9, rho = -0.23, p-value = 1.93e-08.

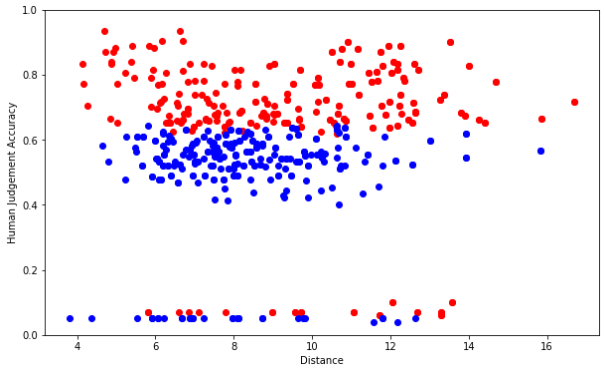


Figure 8

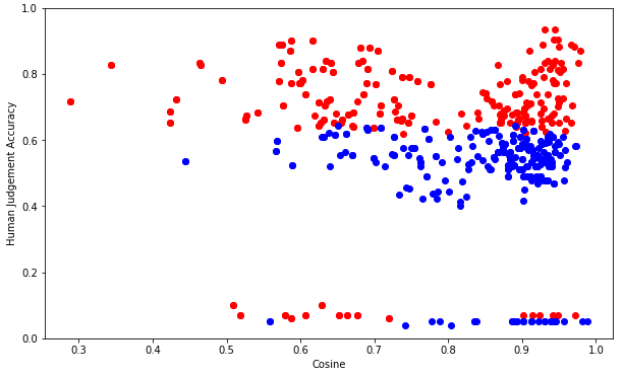


Figure 9

In Figure 10 and Figure 11, the y-axis is the probability of human judging two images in the same category. In Figure 10, the x-axis is the Euclidean distance between two images while in Figure 11, the x-axis is the cosine value between two images. In Figure 10 rho = -0.45, p-value = 1.65e-30 while in Figure 11, rho = 0.42, p-value = 6.26e-27, which shows a strong correlation than accuracy in Figure 8 and Figure 9. The interpretation of this is that if two images are more similar mathematically, either the Euclidean between them is shorter, or in the case that the cosine value is greater, then there's a higher probability that human will judge this image pair as the same category.

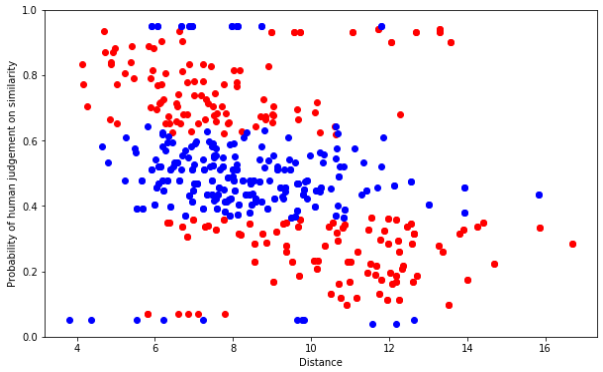


Figure 10

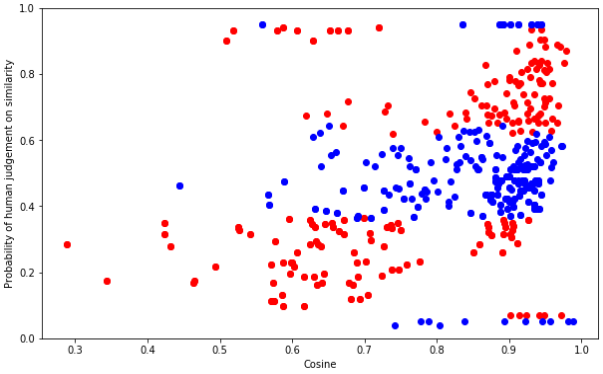


Figure 11

In Figure 5 in the Beta set, the rho and p-value for the correlation between distance and the probability of human judging the same category in the baseline are -0.21 and 0.0038 respectively, the graph is depicted in Figure 12, with x-axis to be distance and y-axis to be the probability. In exp 12 after the whole network, the corresponding rho and p-value

are -0.29 and 5.09e-05, which shows a stronger correlation. The graph in exp 12 is depicted in Figure 13.

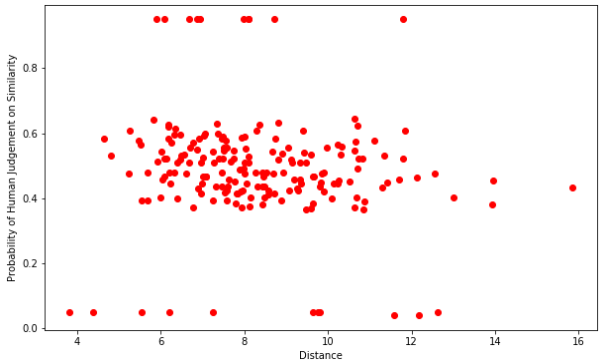


Figure 12

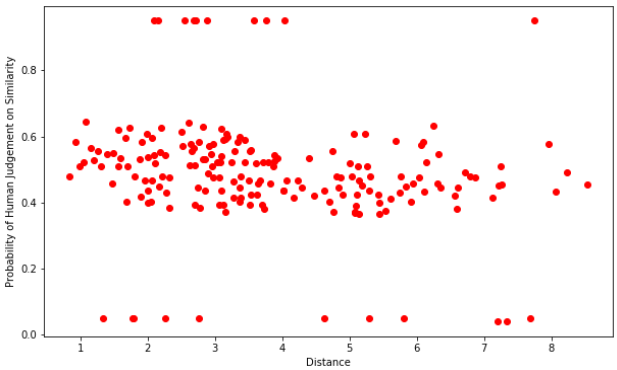


Figure 13

Discussion & Future Work

In the future, I would like to add Convolution, Batch Normalization and Relu in each layer. The reason for choosing this is that Convolution layer can also extract features, Batch Normalization can fasten learning and increase the accuracy. Using Relu because as the autoencoder gets deeper, it tends to have the problem of vanishing gradient and Relu can help with it.

I would like to add Pooling in each layer in the encoder and Unsampling in each layer in the decoder, because Pooling is a way to non-linear downsample the image.