Toward Correlating and Solving Abstract Tasks Using Convolutional Neural Networks Supplementary Material

Kuan-Chuan Peng
Cornell University
kp388@cornell.edu

Tsuhan Chen
Cornell University
tsuhan@cornell.edu

1. Details of Experimental Settings

As we promise in the main paper, Table 1 and the following subsections explain the details of the experimental settings of the three tasks: EMO, AVA, and FAS.

1.1. Emotion Classification (EMO)

All the previous works [16, 18] conduct their experiments in 1-vs-all setting, reporting the average true positive per class of each emotion category using 5-fold cross validation. We follow the same setting, creating the training set and testing set for each emotion category according to the number of the images in each emotion category. Since in 1-vs-all setting, the number of the positive examples is much smaller than that of the negative examples, we repeat the positive examples in the training set such that the number of the positive examples and that of the negative examples are the same. We use the modified training set to avoid that the classification results are biased by the dominating negative examples. In addition to 1-vs-all setting, we also conduct 8-way classification, reporting the average 5-fold cross validation accuracy in Sec. 3.

1.2. Aesthetic Classification (AVA)

We label images with scores $\geq$5 as images with high aesthetic quality and the other images as images with low aesthetic quality. There are 10 training sets and 9 testing sets provided by AVA [17]. For each of the 8 content categories, one training set and one testing set are provided (each with $\sim$2.5k images). AVA [17] also provides one small scale generic training set with 2495 images (training set ID: generic-ss), one large scale generic training set with 19924 images (training set ID: generic-ls), and one generic testing set with 19930 images. In addition, we create a training set with 233012 images (training set ID: generic-vls) such that it includes most images in AVA [17] without overlapping with any of the 9 testing sets.

1.3. Fashion Style Classification (FAS)

There is a score associated with each image in HipsterWars [12]. The higher the score is, the better the image fits the fashion style it is labeled. In [12], the parameter δ is introduced to control the number of images used in training/testing sets. δ varies from 0.1 to 0.5 where $\delta = 0.1$ represents a classification task between the images with top 10% of scores from each style. After the images are selected, we use 9:1 training to testing ratio according to [12].

2. Details of Task Properties and Training Parameters

In the main paper, we apply the convolutional neural network (CNN) proposed in [14] to 8 different abstract tasks with Caffe [9] implementation, training in GPU mode. Since the GPU does not have enough memory to process all the training or testing examples at once, we divide the training and testing set into several batches, which is the solution in Caffe [9] implementation. For each training or testing iteration, we use only one batch of data to update the CNN parameters or to evaluate the testing examples. For all the 8 tasks, we fix the training batch size to 50, and the number of training iterations is chosen such that each training example contributes to updating the CNN parameters at least 20 times. The testing batch size and the number of testing iterations are chosen such that the following two constraints are satisfied:

1. The testing batch size is small enough to be processed by the GPU.

dataset	Artphoto	Painting-91	Painting-91	AVA	HipsterWars	arcDataset	Memorability	Memorability
task	emotion classification	artist classification	artistic style classification	aesthetic classification	fashion style classification	architectural style classification	memorability prediction	interestingness prediction
reference	[16]	[10]	[10]	[17]	[12]	[19]	[8]	dataset: [8]; task: [6]
task ID	EMO	AST	ART	AVA	FAS	ARC	MEM	INT
# classes	8	91	13	2	5	10 / 25	regression task	regression task
# images	806	4266	2338	>250k	1893	2043 / 4786	2222	2222
image type	deviantart [3]	painting	painting	dpchallenge [4]	outfit	architecture	general	general
class labels	fear, sad, etc.	Rubens, Picasso, etc.	Baroque, Cubbism, etc.	high / low aesthetic quality	Bohemian, Goth, etc.	Georgian, Gothic, etc.	memorability	interestingness
# training images	$\sim$645	2275	1250	$\sim$233k	853	300 / 750	1111	1982 / 1000
# testing images	$\sim$160	1991	1088	19930	92	1743 / 4036	1111	240 / 1222
data split	random	specified [10]	specified [10]	specified [17]	random	random	specified [8]	random
# fold(s)	5	1	1	1	100	10	25	10
evaluation metric	1-vs-all accuracy	accuracy	accuracy	accuracy	accuracy	accuracy	ρ	K / ρ
reference of above setting	[18]	[10]	[10]	[17]	[12]	[19]	[8]	[5] / [6]

Table 1. The datasets and associated abstract tasks used in this work along with their properties. We refer to each task by the corresponding task ID listed under each task. The experimental setting for each task is provided at the bottom of the table, where ρ and K are Spearman rank correlation and Kendall tau rank distance between the prediction and the ground truth respectively.

properties \ task ID	AST	ART	ARC	MEM	INT
# of classes	91	13	10 / 25	regression task	regression task
# of training images	2275	1250	200 / 750	1111	1982 / 1000
# of testing images	1991	1088	1743 / 4036	1111	240 / 1222
testing batch size	11	34	21 / 4	11	10 / 26
# of testing iteration	181	32	83 / 1009	101	24 / 47
# of training iteration	7500	3000	150 / 300	500	800 / 400
evaluation metric	accuracy	accuracy	accuracy	ρ	ρ / K

Table 2. The details of task properties and training parameters for AST, ART, ARC, MEM, and INT. For ARC and INT, we list the information of two different experimental settings. ρ and K represent Spearman rank correlation and Kendall tau rank distance respectively.

2. The product of the testing batch size and the number of testing iterations is equal to the number of testing examples, which makes sure that each testing example is evaluated exactly once.

We use the task IDs assigned in the main paper to refer to each task. For AST, ART, ARC, MEM, and INT, the details of task properties and training parameters are summarized in Table 2. For EMO, the corresponding information is listed in Table 3 and Table 4. For AVA and FAS, the details of task properties and training parameters are shown in Table 5 and Table 6 respectively.

3. Experimental Results in Detail

We summarize the experimental results of the selected 8 abstract tasks in Table 8, where the pointers of additional results are also provided if there are multiple experimental settings for that task. From Table 8 to Table 14, we show the results in a coherent way: the **bold** numbers represent the best performance given the specified evaluation metric, and the underlined numbers indicate the performance better than that of "train from scratch." Table 8 shows that in all 8 tasks, at least one of the five training approaches in Table 7 outperforms the state-of-the-art methods. Table 8 also shows that for most of the 8 tasks, the training approaches involving pre-training and fine-tuning usually outperform training from scratch. For AST, ART, and ARC, the complete results are shown in Table 8, where the results of ARC are displayed in the form: 10-way / 25-way classification accuracy and the results of EMO are displayed in the form: 1-vs-all / 8-way classification accuracy. The following paragraphs explain the results of the other tasks in detail.

properties \ emotion category	amusement	anger	awe	contentment	disgust	excitement	fear	sad
fold ID: 1								
# of training images	1122	1148	1126	1184	1174	1148	1096	1018
# of positive testing images	18	7	21	18	13	35	19	31
# of negative testing images	144	155	141	144	149	127	143	131
fold ID: 2								
# of training images	1124	1190	1130	1172	1180	1106	1096	1032
# of positive testing images	18	27	22	11	15	13	18	37
# of negative testing images	143	134	139	150	146	148	143	124
fold ID: 3								
# of training images	1126	1170	1122	1176	1178	1130	1122	1006
# of positive testing images	19	17	18	13	14	25	31	24
# of negative testing images	142	144	143	148	147	136	130	137
fold ID: 4								
# of training images	1138	1166	1126	1176	1184	1112	1098	1030
# of positive testing images	25	15	20	13	17	16	19	36
# of negative testing images	136	146	141	148	144	145	142	125
fold ID: 5								
# of training images	1130	1158	1128	1180	1172	1112	1116	1034
# of positive testing images	21	11	21	15	11	16	28	38
# of negative testing images	140	150	140	146	150	145	133	123

Table 3. The details of task properties and training parameters for EMO in 1-vs-all binary classification setting, where the evaluation metric is average true positive per class. We conduct 5-fold cross validation. The number of the training iterations and the testing batch size are fixed to 500 and 1 respectively.

properties \ fold ID	1	2∼5
# of training images	644	645
# of testing images	162	161
testing batch size	18	23
# of testing iteration	9	7

Table 4. The details of task properties and training parameters for EMO in 8-way classification setting, where the evaluation metric is classification accuracy. We conduct 5-fold cross validation. The number of the training iterations is fixed to 300.

3.1. Fashion Style Classification (FAS)

The complete results are in Table 9, where the numbers are the average 5-way classification accuracy of 100 random training/testing splits. The results of FAS in Table 8 are the ones corresponding to $\delta = 0.5$ in Table 9. From Table 9, we find that for 3 out of 5 different δ values, at least one of the five CNN-based approaches outperforms Kiapour's method [12]. Table 9 also shows that the training approaches involving pre-training and fine-tuning usually outperform training from scratch.

3.2. Memorability Prediction (MEM)

The complete results are in Table 10, where the numbers are the average Spearman rank correlation (ρ) between the prediction and ground truth of 25 specified training/testing splits [8]. For fair comparison, we separate the results without extra information from the ones using extra information in Table 10, where the categories of extra information are also listed. We only report the results of CNN-based training approaches without extra information because the extra information used

properties \ content category	animal	architecture	cityscape	floral	fooddrink	landscape	portrait	stilllife
# of training images	2480	2496	2492	2495	2496	2490	2485	2490
# of testing images	2484	2495	2494	2495	2493	2490	2488	2491
testing batch size	36	5	58	5	9	30	8	53
# of testing iteration	69	499	43	499	277	83	311	47

Table 5. The details of task properties and training parameters for AVA in binary classification setting, where the evaluation metric is classification accuracy. The number of the training iterations for content-based training sets is fixed to 1000. For the three different generic training sets–generic-ss, generic-ls, and generic-vls, the numbers of the training iterations are set to 1000, 8000, and 100000 respectively. For the generic testing set containing 19930 images, the testing batch size and the number of the testing iterations are set to 10 and 1993 respectively.

properties \ δ	0.1	0.2	0.3	0.4	0.5
# of training images	171	342	512	680	853
# of testing images	16	35	54	74	92
testing batch size	16	35	54	37	46
# of testing iteration	1	1	1	2	2
# of training iteration	100	150	250	300	350

Table 6. The details of task properties and training parameters for FAS in 5-way classification setting, where the evaluation metric is classification accuracy. δ is introduced in [12] to control the number of images used in training/testing sets. For example, $\delta = 0.1$ represents a classification task between the images with top 10% of scores from each style.

training approach ID	description
pt ImageNet + ft	Pre-train with M_{ImageNet} and fine-tune all the CNN parameters using the training set.
pt ImageNet + ft-fc8	The same as "pt ImageNet + ft" except that only the CNN parameters associated with the edges directly connected to the output layer are allowed to be updated using the training set.
pt AVA + ft	Pre-train with M_{AVA} and fine-tune all the CNN parameters using the training set.
pt AVA + ft-fc8	The same as "pt AVA + ft" except that only the CNN parameters associated with the edges directly connected to the output layer are allowed to be updated using the training set.
train from scratch	Randomly initialize all the CNN parameters and train with the training set.

Table 7. The five different CNN training approaches used in this work. M_{ImageNet} is the Caffe [9] reference model trained for ImageNet [2] classification, and M_{AVA} is our trained reference model for AVA [17] classification. In this work, we refer to each training method by its training approach ID.

task ID evaluation metric	EMO 1-vs-all / 8-way classification accuracy (%)	AST accuracy (%)	ART accuracy (%)	AVA accuracy (%)	FAS accuracy (%)	ARC accuracy (%)	MEM ρ	INT ρ
previous work	<u>63.163</u> [18] / n / a	<u>53.100</u> [10]	<u>62.200</u> [10]	73.250 [15]	<u>70.971</u> [12]	<u>69.170</u> / <u>46.210</u> [19]	<u>0.500</u> [11]	<u>0.600</u> [6]
pt ImageNet + ft	60.127 / <u>33.123</u>	<u>**56.102**</u>	<u>**68.290**</u>	n / a	<u>**71.294**</u>	<u>**71.159**</u> / <u>**52.953**</u>	<u>**0.520**</u>	<u>**0.643**</u>
pt ImageNet + ft-fc8	<u>**64.724**</u> / <u>**34.116**</u>	<u>53.541</u>	<u>65.165</u>	n / a	<u>66.228</u>	<u>67.246</u> / <u>51.469</u>	-0.140	0.339
pt AVA + ft	59.836 / <u>27.670</u>	<u>25.615</u>	<u>40.625</u>	n / a	<u>57.337</u>	<u>35.841</u> / <u>20.401</u>	0.368	<u>0.511</u>
pt AVA + ft-fc8	60.644 / 22.090	4.671	18.015	n / a	27.554	18.233 / 8.290	0.080	-0.113
train from scratch	61.572 / 22.712	21.698	38.327	**74.436**	54.304	21.532 / 12.386	0.372	0.382
additional results	Table 12	n / a	n / a	Table 13, 14	Table 9	n / a	Table 10	Table 11

Table 8. The summary of the experimental results of the 8 abstract tasks listed in Table 1 using the five training approaches in Table 7. In this table, ρ is the Spearman rank correlation between the prediction and the ground truth. The **bold** numbers represent the best performance given the specified evaluation metric, and the <u>underlined</u> numbers indicate the performance better than that of "train from scratch". The pointers of the additional results are also provided if there are multiple experimental settings for the task. This table shows that CNN-based approaches outperform the state-of-the-art methods on all 8 tasks. For most of the 8 tasks, the training approaches involving pre-training and fine-tuning usually outperform training from scratch.

δ	0.1	0.2	0.3	0.4	0.5
# training images	171	342	512	680	853
# testing images	16	35	54	74	92
Kiapour's method [12]	<u>70.194</u>	<u>**75.728**</u>	<u>**75.243**</u>	<u>72.330</u>	<u>70.971</u>
pt ImageNet + ft	<u>**73.438**</u>	<u>72.486</u>	<u>72.019</u>	<u>**74.500**</u>	<u>**71.294**</u>
pt ImageNet + ft-fc8	<u>68.875</u>	<u>67.943</u>	<u>68.148</u>	<u>67.649</u>	<u>66.228</u>
pt AVA + ft	47.563	<u>49.514</u>	<u>53.685</u>	<u>56.797</u>	<u>57.337</u>
pt AVA + ft-fc8	24.875	27.257	25.759	27.622	27.554
train from scratch	49.188	43.429	48.093	51.987	54.304

Table 9. The experimental results of FAS (described in Table 1) in terms of the average 5-way classification accuracy (%) of 100 random training/testing splits. δ is introduced in [12] to control the number of images used in training/testing sets. For example, $\delta = 0.1$ represents a classification task between the images with top 10% of scores from each style. We use the training approach ID in Table 7 to refer to each training method. The **bold** numbers represent the best accuracy under each δ, and the <u>underlined</u> numbers indicate the accuracy better than that of "train from scratch" under the corresponding δ.

ρ	w/o extra info	w/ extra info	extra info
Isola's method [7]	<u>0.460</u>	0.540	object, scene visual attribute
Celikkale's method [1]	<u>0.470</u>	0.470	object, saliency
Khosla's method [11]	<u>0.500</u>	n / a	n / a
Kim's method [13]	n / a	**0.580**	object, scene visual attribute
human [8]	<u>*0.750*</u>	n / a	n / a
pt ImageNet + ft	<u>**0.520**</u>	n / a	n / a
pt ImageNet + ft-fc8	-0.140	n / a	n / a
pt AVA + ft	0.368	n / a	n / a
pt AVA + ft-fc8	0.080	n / a	n / a
train from scratch	0.372	n / a	n / a

Table 10. The experimental results of MEM (described in Table 1) in terms of the average Spearman rank correlation (ρ) between the prediction and ground truth of 25 specified training/testing splits [8]. For fair comparison, we separate the results without extra information from those using extra information (the used extra information is also provided). We only apply the training methods in Table 7 without using extra information because the extra information used in each previous work is inconsistent. The **bold** numbers represent the best ρ out of all the non-human methods, and the <u>underlined</u> numbers indicate the ρ better than that of "train from scratch."

in each previous work is inconsistent. The results of MEM in Table 8 are the partial results in the column "w/o extra info" of Table 10, where we find that the performance of "pt ImageNet + ft" is the best out of all the non-human methods in that column.

3.3. Interestingness Prediction (INT)

The complete results are in Table 11, where the numbers are the average Spearman rank correlation (ρ) or Kendall tau rank distance (K) between the prediction and ground truth of 10 random training/testing splits. For ρ, higher is better. For K, lower is better. The results of INT in Table 8 are the results using the evaluation metric ρ in Table 11, where we find that at least one of the five CNN-based approaches outperforms Gygli's method [6] under the evaluation metric ρ. However, CNN-based approaches perform worse than Fu's method [5] under the evaluation metric K. Even Fu et al. [5] show that their method is better than Gygli's method [6] in terms of K, we suspect that Fu et al. [5] happen to find an evaluation metric (K) in favor of their method without reporting the results using ρ, the evaluation metric used in the first work [6] introducing the task INT.

3.4. Emotion Classification (EMO)

Table 8 shows the average 5-fold cross validation accuracy in both 1-vs-all and 8-way classification settings. In fact, all the previous works [16, 18] related to the task EMO report their results in average true positive per class of each emotion category using 5-fold cross validation under 1-vs-all setting. Therefore, we follow the experimental setting of [16, 18], reporting the detail results in Table 12, where we find that in 7 out of 8 emotion categories, at least one of the five CNN-based approaches

evaluation metric	ρ	K
previous work	<u>0.600</u> [6]	**<u>0.110</u>** [5]
pt ImageNet + ft	**<u>0.643</u>**	<u>0.246</u>
pt ImageNet + ft-fc8	0.339	0.377
pt AVA + ft	<u>0.511</u>	<u>0.354</u>
pt AVA + ft-fc8	-0.113	0.525
train from scratch	0.382	0.361

Table 11. The experimental results of INT (described in Table 1) in terms of the average Spearman rank correlation (ρ) or Kendall tau rank distance (K) between the prediction and ground truth of 10 random training/testing splits. For ρ, higher is better. For K, lower is better. We use the training approach ID in Table 7 to refer to each training method. The **bold** numbers represent the best performance under the specified evaluation metric, and the <u>underlined</u> numbers indicate the performance better than that of "train from scratch" under the specified evaluation metric. Although CNN-based approaches perform worse than Fu's method [5] under K, "pt ImageNet + ft" outperforms Gygli's method [6] under ρ, the evaluation metric used in the first work [6] introducing the task INT.

emotion category	amusement	anger	awe	contentment	disgust	excitement	fear	sad
# images	101	77	102	70	70	105	115	166
Wang's method [18]	**<u>70.000</u>**	<u>57.000</u>	<u>59.000</u>	<u>64.800</u>	58.500	<u>67.000</u>	68.000	61.000
pt ImageNet + ft	61.982	<u>56.197</u>	<u>56.623</u>	58.599	54.968	<u>67.264</u>	64.915	60.466
pt ImageNet + ft-fc8	<u>69.412</u>	<u>63.925</u>	<u>61.956</u>	**<u>66.092</u>**	59.864	**<u>67.365</u>**	67.999	61.177
pt AVA + ft	<u>65.500</u>	<u>53.941</u>	<u>58.091</u>	60.819	55.445	58.937	66.877	59.082
pt AVA + ft-fc8	56.352	**<u>69.553</u>**	**<u>66.954</u>**	62.186	57.649	58.291	58.397	55.774
train from scratch	63.266	53.216	56.254	62.458	**61.635**	64.116	**70.155**	**61.476**

Table 12. The experimental results of EMO (described in Table 1) in terms of the average true positive per class (%) using 5-fold cross validation under 1-vs-all setting. The number of images in each emotion category is also shown. We use the training approach ID in Table 7 to refer to each training method. The **bold** numbers represent the best performance under each emotion category, and the <u>underlined</u> numbers indicate the performance better than that of "train from scratch" under each emotion category. In all the 8 categories except "amusement", CNN-based approach performs the best out of all the listed methods.

outperforms Wang's method [18]. The reported results in Table 8 under 1-vs-all setting are the average performances across all the 8 emotion categories in Table 12.

3.5. Aesthetic Classification (AVA)

Murray et al. [17] train on three different kinds of training sets (generic-ss, content-based images, and generic-ls) separately and test on content-based testing sets. We follow their experimental setting, comparing our results with theirs in Table 13, where the numbers represent the binary classification accuracy. We separate the results into three sections according to the training set. In addition, we also train on the the training set "generic-vls" and report the accuracy. Table 13 shows that regardless of the training set used, "pt AVA + ft-fc8" performs the best out of all the listed methods, which is not surprising because the dataset used for pre-training in "pt AVA + ft-fc8" fits the task AVA well. With great initialization of CNN parameters, "pt AVA + ft-fc8" outperforms all the other methods listed in Table 13.

In contrast, the performance of the training methods involving "pt ImageNet" is even worse than that of "train from scratch," which is surprising given that most CNN-related works report the positive effect brought by "pt ImageNet." In fact, AVA is the only abstract task in Table 1 such that pre-training from ImageNet [2] performs worse than "train from scratch," which may result from the dissimilarity between the class labels in ImageNet [2] and those in AVA [17]. The fact that the training approaches involving "pt AVA" outperform those involving "pt ImageNet" supports that we should select the dataset more relevant to the task of interest for pre-training.

Murray et al. [17] also make the following two claims based on their results: 1: The performance of training from the content-based training set is better than that of training from "generic-ss." 2: The performance of training from "generic-ls" is better than that of training from the content-based training set. Their first claim is consistent with our results in Table 13. However, their second claim is inconsistent with the results of "pt ImageNet + ft-fc8" and "pt AVA + ft-fc8," which means that training with a large number of generic images may be worse than training with a small number of images with relevant contents.

In addition, we train on all the 10 training sets provided in AVA [17] and the training set "generic-vls" separately, and test

testing set category	animal	architecture	cityscape	floral	fooddrink	landscape	portrait	stilllife
training set	2495 generic images (training set ID: generic-ss)							
Murray's method (color) [17]	61.778	64.000	63.704	62.963	62.519	65.926	58.963	62.667
Murray's method (SIFT) [17]	62.222	64.889	62.667	65.185	62.963	65.630	58.815	64.889
pt ImageNet + ft	69.404	69.379	70.249	68.497	63.979	74.859	69.976	60.699
pt ImageNet + ft-fc8	62.359	68.257	67.322	63.968	61.532	73.414	63.344	58.932
pt AVA + ft	72.142	73.587	71.852	69.940	69.715	77.590	75.804	66.038
pt AVA + ft-fc8	**74.638**	**76.794**	**74.379**	**72.866**	**72.764**	**79.478**	**77.733**	**67.724**
train from scratch	70.934	74.229	71.371	70.381	67.670	77.390	75.924	62.626
training set	∼2.5k content-based images							
Murray's method (color) [17]	64.000	65.185	66.667	66.074	65.037	67.556	62.370	63.704
Murray's method (SIFT) [17]	65.926	64.741	65.926	67.852	65.185	68.889	62.667	66.222
pt ImageNet + ft	70.491	72.425	71.572	69.900	65.503	77.390	74.196	63.990
pt ImageNet + ft-fc8	69.968	70.541	72.494	64.209	64.380	78.755	74.920	65.155
pt AVA + ft	72.705	75.391	72.294	70.261	68.311	76.145	76.688	64.071
pt AVA + ft-fc8	**74.718**	**76.433**	**74.579**	**72.385**	**72.483**	**79.759**	**78.256**	**67.965**
train from scratch	70.934	74.228	70.730	70.381	69.114	77.390	75.965	62.625
training set	19924 generic images (training set ID: generic-ls)							
Murray's method (color) [17]	64.741	68.741	68.889	67.259	66.370	69.185	62.963	66.667
Murray's method (SIFT) [17]	64.889	66.963	68.000	67.556	65.926	70.074	63.704	65.630
pt ImageNet + ft	72.061	74.188	73.376	68.657	67.629	76.546	73.674	64.593
pt ImageNet + ft-fc8	65.419	60.160	60.666	67.455	63.418	61.687	66.801	61.381
pt AVA + ft	72.987	75.391	72.534	72.305	69.956	78.514	76.769	64.593
pt AVA + ft-fc8	**74.517**	**76.914**	**74.459**	**72.946**	**72.603**	**79.398**	**77.974**	**67.764**
train from scratch	70.934	74.229	71.371	70.381	67.670	77.390	75.924	62.626
training set	233012 generic images (training set ID: generic-vls)							
train from scratch	74.597	76.713	74.058	73.307	71.240	79.960	78.296	67.322

Table 13. The experimental results of AVA (described in Table 1) in terms of the binary classification accuracy (%). We train on four different kinds of training sets and test on the content-based testing set. We use the training approach ID in Table 7 to refer to each training method. The **bold** numbers represent the best accuracy under each content category, and the underlined numbers indicate the accuracy better than that of "train from scratch" under each content category. In general, "pt AVA + ft-fc8" performs the best out of all the listed methods, but the training methods involving "pt ImageNet" perform even worse than "train from scratch."

training set category	animal	architecture	cityscape	floral	fooddrink	landscape	portrait	stilllife	generic-ss	generic-ls	generic-vls
testing set	19930 generic images										
Lu's method [15]	n / a	n / a	n / a	n / a	n / a	n / a	n / a	n / a	n / a	n / a	73.250
pt ImageNet + ft	65.218	65.519	65.153	64.782	61.470	64.822	66.428	65.434	68.379	71.340	n / a
pt ImageNet + ft-fc8	66.864	64.285	63.899	58.249	61.651	68.635	68.254	66.172	64.972	63.327	n / a
pt AVA + ft	71.706	71.109	69.734	71.039	68.083	70.582	71.400	65.690	72.133	72.875	n / a
pt AVA + ft-fc8	**74.591**	**74.676**	**74.671**	**74.516**	**74.365**	**74.501**	**74.671**	**74.305**	**74.636**	**74.611**	n / a
train from scratch	71.315	71.315	65.720	71.295	68.610	71.259	71.249	71.315	71.315	71.315	**74.436**

Table 14. The experimental results of AVA (described in Table 1) in terms of the binary classification accuracy (%). We train on all 10 training sets provided by AVA [17] and the training set "generic-vls", and test on the generic testing set. We use the training approach ID in Table 7 to refer to each training method. The **bold** numbers represent the best accuracy under each training set, and the underlined numbers indicate the accuracy better than that of "train from scratch" under each training set. In general, "pt AVA + ft-fc8" performs the best out of all the listed methods, but the training methods involving "pt ImageNet" perform even worse than "train from scratch."

on the generic testing set. The accuracy is reported in Table 14, where "pt AVA + ft-fc8" performs the best out of all the listed methods. In the rightmost column of Table 14, we show that our result is better than that of Lu's method [15] under the same experimental setting, which is also the results of AVA shown in Table 8. The best performance in Lu's work [15] (74.460%)

learned CNN features	F_ART	F_AST	F_CAL	F_ARC	F_EMO	F_AVA	F_FAS	F_MEM	F_INT
feature index	1	2	3	4	5	6	7	8	9

Table 15. The feature index we assign to each leaned CNN features to specify the concatenated features in Table 16 to Table 24. For example, "F_258" means concatenating the three features: F_AST (feature index: 2), F_EMO (feature index: 5), and F_MEM (feature index: 8).

is close to our result (74.436%), but their best performance is achieved by using the extra information from the style labels which we do not use. The reason why our method outperforms Lu's method [15] is that we use more layers and nodes in our CNN architecture compared to theirs.

4. Complete Results of Concatenating the Learned CNN Features

As we promise in the main paper, in Table 16 to Table 24, we provide the performances of all the 264 feature concatenating settings in each of the 9 tasks, including the 8 abstract tasks listed in Table 1 and the Caltech-101 object classification task (we use CAL as its task ID). We use the notation "F_T" to represent the CNN features learned from the task T. To specify the concatenated features in Table 16 to Table 24, we use the feature index we assign to each learned CNN features (Table 15). For example, "F_25" means concatenating F_AST (feature index: 2) and F_EMO (feature index: 5). Table 16 to Table 24 support the following claims:

1. Concatenating the CNN features learned from other tasks can improve the performance in each task.
2. Concatenating all the features (F_123456789) never performs the best in all the 9 tasks under experiment, which suggests that we should concatenate useful features instead of all the features.

References

[1] B. Celikkale, A. Erdem, and E. Erdem. Visual attention-driven spatial pooling for image memorability. In *CVPRW*, pages 976–983, 2013.

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and F.-F. Li. Imagenet: A large-scale hierarchical image database. In *CVPR*, pages 248–255, 2009.

[3] deviantart. `http://www.deviantart.com`.

[4] dpchallenge. `http://www.dpchallenge.com/`.

[5] Y. Fu, T. M. Hospedales, T. Xiang, S. Gong, and Y. Yao. Interestingness prediction by robust learning to rank. In *ECCV*, pages 488–503, 2014.

[6] M. Gygli, H. Grabner, H. Riemenschneider, F. Nater, and L. V. Gool. The interestingness of images. In *ICCV*, pages 1633–1640, 2013.

[7] P. Isola, J. Xiao, D. Parikh, A. Torralba, and A. Oliva. What makes a photograph memorable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1469–1482, 2014.

[8] P. Isola, J. Xiao, A. Torralba, and A. Oliva. What makes an image memorable? In *CVPR*, pages 145–152, 2011.

[9] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.

[10] F. S. Khan, S. Beigpour, J. V. D. Weijer, and M. Felsberg. Painting-91: a large scale database for computational painting categorization. *Machine Vision and Applications*, 25:1385–1397, 2014.

[11] A. Khosla, J. Xiao, A. Torralba, and A. Oliva. Memorability of image regions. In *NIPS*, pages 296–304, 2012.

[12] M. H. Kiapour, K. Yamaguchi, A. C. Berg, and T. L. Berg. Hipster wars: Discovering elements of fashion styles. In *ECCV*, pages 472–488, 2014.

[13] J. Kim, S. Yoon, and V. Pavlovic. Relative spatial features for image memorability. In *ACMMM*, pages 761–764, 2013.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, pages 1097–1105. 2012.

[15] X. Lu, Z. Lin, H. Jin, J. Yang, and J. Z. Wang. RAPID: rating pictorial aesthetics using deep learning. In *ACMMM*, 2014.

[16] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In *Proceedings of the International Conference on Multimedia*, pages 83–92, 2010.

[17] N. Murray, L. Marchesotti, and F. Perronnin. AVA: A large-scale database for aesthetic visual analysis. In *CVPR*, pages 2408–2415, 2012.

[18] X. Wang, J. Jia, J. Yin, and L. Cai. Interpretable aesthetic features for affective image classification. In *IEEE International Conference on Image Processing*, pages 3230–3234, 2013.

[19] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi. Architectural style classification using multinomial latent logistic regression. In *ECCV*, pages 600–615, 2014.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	67.555	F_1259	69.485	F_12458	69.945	F_123469	69.853	F_1234789	69.577
F_2	67.555	F_1267	70.221	F_12459	69.761	F_123478	70.129	F_1235678	70.496
F_3	61.121	F_1268	70.404	F_12467	69.853	F_123479	69.853	F_1235679	69.669
F_4	60.938	F_1269	69.669	F_12468	69.577	F_123489	69.577	F_1235689	69.945
F_5	60.938	F_1278	70.588	F_12469	69.853	F_123567	70.037	F_1235789	69.761
F_6	56.985	F_1279	69.302	F_12478	69.302	F_123568	71.048	F_1236789	69.485
F_7	57.813	F_1289	68.842	F_12479	69.669	F_123569	69.761	F_1245678	69.945
F_8	51.838	F_1345	68.107	F_12489	69.393	F_123578	70.680	F_1245679	69.485
F_9	55.790	F_1346	68.934	F_12567	69.669	F_123579	69.945	F_1245689	69.945
F_12	70.129	F_1347	69.302	F_12568	70.588	F_123589	69.577	F_1245789	69.669
F_13	68.658	F_1348	68.382	F_12569	69.761	F_123678	70.772	F_1246789	69.485
F_14	68.658	F_1349	68.015	F_12578	70.404	F_123679	70.221	F_1256789	70.313
F_15	68.382	F_1356	69.210	F_12579	69.853	F_123689	69.761	F_1345678	68.842
F_16	67.739	F_1357	69.945	F_12589	70.404	F_123789	69.485	F_1345679	68.290
F_17	67.647	F_1358	68.842	F_12678	70.864	F_124567	69.118	F_1345689	68.199
F_18	67.831	F_1359	68.566	F_12679	69.485	F_124568	69.853	F_1345789	68.658
F_19	67.096	F_1367	69.761	F_12689	69.118	F_124569	70.037	F_1346789	68.290
F_123	69.853	F_1368	69.853	F_12789	69.302	F_124578	69.669	F_1356789	68.842
F_124	69.302	F_1369	67.923	F_13456	68.107	F_124579	69.669	F_1456789	68.658
F_125	69.393	F_1378	68.842	F_13457	68.658	F_124589	69.945	F_12345678	70.129
F_126	69.945	F_1379	68.474	F_13458	68.842	F_124678	69.577	F_12345679	69.761
F_127	70.129	F_1389	68.199	F_13459	68.382	F_124679	69.577	F_12345689	69.026
F_128	70.129	F_1456	67.739	F_13467	69.118	F_124689	69.761	F_12345789	69.485
F_129	69.302	F_1457	67.923	F_13468	68.750	F_124789	69.485	F_12346789	69.393
F_134	69.026	F_1458	68.474	F_13469	68.107	F_125678	70.404	F_12356789	69.669
F_135	69.393	F_1459	67.555	F_13478	68.842	F_125679	69.853	F_12456789	69.945
F_136	68.750	F_1467	68.658	F_13479	68.658	F_125689	70.588	F_13456789	68.566
F_137	69.302	F_1468	68.750	F_13489	68.107	F_125789	70.496	F_123456789	69.210
F_138	69.485	F_1469	68.199	F_13567	69.485	F_126789	69.118		
F_139	68.015	F_1478	68.750	F_13568	68.934	F_134567	68.842		
F_145	68.107	F_1479	68.290	F_13569	68.474	F_134568	68.934		
F_146	68.382	F_1489	67.831	F_13578	69.026	F_134569	68.658		
F_147	68.842	F_1567	68.750	F_13579	68.750	F_134578	68.566		
F_148	68.934	F_1568	68.934	F_13589	68.934	F_134579	68.199		
F_149	67.923	F_1569	68.474	F_13678	69.302	F_134589	68.750		
F_156	68.474	F_1578	68.750	F_13679	68.290	F_134678	68.842		
F_157	68.750	F_1579	68.658	F_13689	68.290	F_134679	68.658		
F_158	68.199	F_1589	67.555	F_13789	68.107	F_134689	68.107		
F_159	68.474	F_1678	68.566	F_14567	67.923	F_134789	68.474		
F_167	67.555	F_1679	68.290	F_14568	68.750	F_135678	69.026		
F_168	68.750	F_1689	67.004	F_14569	67.647	F_135679	68.658		
F_169	67.279	F_1789	67.555	F_14578	68.934	F_135689	69.026		
F_178	67.923	F_12345	69.669	F_14579	68.750	F_135789	68.842		
F_179	68.566	F_12346	69.485	F_14589	68.107	F_136789	68.474		
F_189	66.912	F_12347	69.853	F_14678	69.210	F_145678	69.302		
F_1234	69.485	F_12348	69.945	F_14679	67.831	F_145679	68.934		
F_1235	69.485	F_12349	70.037	F_14689	68.382	F_145689	68.290		
F_1236	69.577	F_12356	69.761	F_14789	68.199	F_145789	68.199		
F_1237	70.129	F_12357	69.853	F_15678	69.302	F_146789	68.566		
F_1238	70.129	F_12358	70.864	F_15679	68.842	F_156789	68.015		
F_1239	69.945	F_12359	69.577	F_15689	67.463	F_1234567	69.669		
F_1245	68.382	F_12367	70.037	F_15789	67.371	F_1234568	70.496		
F_1246	69.118	F_12368	70.680	F_16789	67.739	F_1234569	69.669		
F_1247	69.669	F_12369	69.853	F_123456	69.853	F_1234578	70.221		
F_1248	69.485	F_12378	70.680	F_123457	69.577	F_1234579	69.853		
F_1249	69.945	F_12379	70.404	F_123458	70.496	F_1234589	69.393		
F_1256	69.393	F_12389	69.669	F_123459	69.669	F_1234678	70.221		
F_1257	69.761	F_12456	68.658	F_123467	69.118	F_1234679	69.761		
F_1258	70.496	F_12457	68.842	F_123468	70.496	F_1234689	69.761		

Table 16. The complete results of concatenating the learned CNN features in the task ART. This table shows that concatenating the CNN features learned from other tasks can improve the performance of ART and that concatenating all the features does not achieve the best performance.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	48.569	F_1267	56.454	F_12467	56.103	F_123489	54.194	F_1235789	54.294
F_2	55.148	F_1268	55.902	F_12468	55.299	F_123567	56.605	F_1236789	53.893
F_3	45.907	F_1269	54.395	F_12469	54.746	F_123568	55.902	F_1245678	55.701
F_4	46.760	F_1278	56.454	F_12478	55.801	F_123569	54.797	F_1245679	54.897
F_5	47.564	F_1279	54.395	F_12479	54.847	F_123578	56.354	F_1245689	54.043
F_6	41.487	F_1289	53.441	F_12489	53.993	F_123579	54.998	F_1245789	54.395
F_7	44.751	F_2345	56.103	F_12567	56.705	F_123589	53.993	F_1246789	54.194
F_8	37.167	F_2346	55.450	F_12568	56.052	F_123678	55.902	F_1256789	53.692
F_9	40.532	F_2347	56.052	F_12569	54.546	F_123679	54.546	F_2345678	55.299
F_12	56.404	F_2348	54.897	F_12578	56.354	F_123689	54.043	F_2345679	54.746
F_23	56.605	F_2349	54.093	F_12579	54.847	F_123789	53.943	F_2345689	53.842
F_24	55.098	F_2356	57.258	F_12589	53.893	F_124567	55.801	F_2345789	53.893
F_25	57.308	F_2357	57.107	F_12678	56.454	F_124568	55.650	F_2346789	53.893
F_26	55.349	F_2358	55.299	F_12679	54.596	F_124569	54.646	F_2356789	53.943
F_27	56.103	F_2359	54.194	F_12689	53.340	F_124578	55.650	F_2456789	53.893
F_28	55.550	F_2367	56.555	F_12789	53.742	F_124579	54.746	F_12345678	55.500
F_29	53.541	F_2368	54.847	F_23456	56.103	F_124589	53.993	F_12345679	54.897
F_123	55.952	F_2369	54.194	F_23457	56.253	F_124678	55.902	F_12345689	54.445
F_124	55.952	F_2378	55.349	F_23458	54.847	F_124679	54.746	F_12345789	54.596
F_125	56.454	F_2379	54.596	F_23459	54.445	F_124689	53.993	F_12346789	54.445
F_126	56.705	F_2389	53.390	F_23467	56.052	F_124789	54.194	F_12356789	54.445
F_127	56.253	F_2456	55.902	F_23468	54.947	F_125678	56.555	F_12456789	54.244
F_128	55.902	F_2457	55.801	F_23469	54.043	F_125679	54.897	F_23456789	54.144
F_129	54.395	F_2458	55.851	F_23478	55.098	F_125689	53.792	F_123456789	54.596
F_234	55.299	F_2459	54.144	F_23479	54.345	F_125789	54.043		
F_235	56.806	F_2467	55.450	F_23489	53.943	F_126789	53.692		
F_236	56.404	F_2468	55.701	F_23567	56.956	F_234567	56.052		
F_237	56.404	F_2469	53.893	F_23568	55.450	F_234568	54.696		
F_238	54.746	F_2478	55.650	F_23569	54.043	F_234569	54.445		
F_239	54.194	F_2479	54.244	F_23578	55.952	F_234578	55.249		
F_245	55.902	F_2489	53.591	F_23579	54.345	F_234579	54.546		
F_246	55.098	F_2567	57.459	F_23589	53.641	F_234589	54.043		
F_247	55.751	F_2568	56.203	F_23678	55.650	F_234678	55.450		
F_248	55.299	F_2569	54.244	F_23679	54.546	F_234679	54.445		
F_249	53.792	F_2578	56.504	F_23689	53.491	F_234689	53.692		
F_256	57.107	F_2579	53.993	F_23789	53.842	F_234789	53.742		
F_257	57.509	F_2589	53.290	F_24567	55.952	F_235678	55.902		
F_258	56.303	F_2678	55.701	F_24568	55.751	F_235679	54.294		
F_259	53.792	F_2679	54.194	F_24569	54.244	F_235689	53.692		
F_267	56.605	F_2689	53.340	F_24578	55.801	F_235789	53.943		
F_268	55.349	F_2789	53.591	F_24579	54.445	F_236789	53.742		
F_269	53.290	F_12345	55.902	F_24589	53.943	F_245678	56.103		
F_278	55.701	F_12346	55.801	F_24678	55.349	F_245679	54.646		
F_279	53.792	F_12347	55.500	F_24679	54.093	F_245689	53.591		
F_289	53.441	F_12348	55.098	F_24689	53.842	F_245789	53.792		
F_1234	55.600	F_12349	54.546	F_24789	53.541	F_246789	53.641		
F_1235	56.454	F_12356	56.555	F_25678	56.705	F_256789	53.742		
F_1236	56.153	F_12357	56.605	F_25679	54.244	F_1234567	56.203		
F_1237	56.806	F_12358	55.952	F_25689	53.591	F_1234568	55.399		
F_1238	55.751	F_12359	54.947	F_25789	53.792	F_1234569	54.696		
F_1239	54.445	F_12367	56.655	F_26789	53.541	F_1234578	55.600		
F_1245	55.851	F_12368	55.851	F_123456	55.801	F_1234579	54.998		
F_1246	55.851	F_12369	54.495	F_123457	56.354	F_1234589	54.546		
F_1247	56.002	F_12378	56.103	F_123458	55.349	F_1234678	55.399		
F_1248	55.650	F_12379	54.596	F_123459	54.596	F_1234679	54.947		
F_1249	54.495	F_12389	54.093	F_123467	55.600	F_1234689	54.144		
F_1256	56.906	F_12456	55.801	F_123468	55.299	F_1234789	54.546		
F_1257	56.806	F_12457	55.851	F_123469	54.546	F_1235678	56.655		
F_1258	55.952	F_12458	55.198	F_123478	55.550	F_1235679	54.947		
F_1259	54.696	F_12459	54.495	F_123479	54.847	F_1235689	54.194		

Table 17. The complete results of concatenating the learned CNN features in the task AST. This table shows that concatenating the CNN features learned from other tasks can improve the performance of AST and that concatenating all the features does not achieve the best performance.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	81.080	F_1367	88.394	F_13467	88.224	F_123489	85.949	F_1235789	85.745
F_2	81.066	F_1368	87.267	F_13468	87.538	F_123567	87.633	F_1236789	85.691
F_3	88.217	F_1369	86.146	F_13469	86.418	F_123568	87.280	F_1345678	87.430
F_4	83.667	F_1378	87.246	F_13478	87.518	F_123569	86.207	F_1345679	86.554
F_5	70.771	F_1379	86.268	F_13479	86.594	F_123578	87.246	F_1345689	85.868
F_6	80.475	F_1389	85.426	F_13489	85.766	F_123579	86.357	F_1345789	86.024
F_7	81.514	F_2345	87.823	F_13567	87.986	F_123589	85.623	F_1346789	85.942
F_8	65.643	F_2346	88.020	F_13568	87.178	F_123678	87.273	F_1356789	85.677
F_9	74.316	F_2347	88.020	F_13569	86.180	F_123679	86.309	F_2345678	87.423
F_13	88.231	F_2348	87.328	F_13578	87.226	F_123689	85.589	F_2345679	86.587
F_23	87.823	F_2349	86.431	F_13579	86.336	F_123789	85.650	F_2345689	86.037
F_34	88.244	F_2356	87.769	F_13589	85.535	F_134567	87.980	F_2345789	86.065
F_35	87.728	F_2357	87.749	F_13678	87.409	F_134568	87.375	F_2346789	85.976
F_36	88.319	F_2358	87.124	F_13679	86.323	F_134569	86.492	F_2356789	85.773
F_37	88.299	F_2359	86.160	F_13689	85.487	F_134578	87.334	F_3456789	85.942
F_38	86.975	F_2367	88.007	F_13789	85.677	F_134579	86.526	F_12345678	87.402
F_39	85.834	F_2368	87.307	F_23456	87.864	F_134589	85.949	F_12345679	86.526
F_123	87.742	F_2369	86.227	F_23457	87.803	F_134678	87.470	F_12345689	85.881
F_134	88.136	F_2378	87.321	F_23458	87.294	F_134679	86.608	F_12345789	85.997
F_135	87.980	F_2379	86.295	F_23459	86.486	F_134689	85.847	F_12346789	85.963
F_136	88.340	F_2389	85.474	F_23467	87.973	F_134789	85.922	F_12356789	85.711
F_137	88.292	F_3456	87.993	F_23468	87.477	F_135678	87.226	F_13456789	86.003
F_138	87.253	F_3457	88.020	F_23469	86.424	F_135679	86.391	F_23456789	86.051
F_139	86.065	F_3458	87.158	F_23478	87.491	F_135689	85.562	F_123456789	85.969
F_234	88.027	F_3459	86.200	F_23479	86.404	F_135789	85.643		
F_235	87.633	F_3467	88.299	F_23489	85.922	F_136789	85.698		
F_236	87.932	F_3468	87.559	F_23567	87.735	F_234567	87.844		
F_237	87.939	F_3469	86.289	F_23568	87.328	F_234568	87.362		
F_238	87.131	F_3478	87.572	F_23569	86.214	F_234569	86.540		
F_239	86.119	F_3479	86.526	F_23578	87.328	F_234578	87.368		
F_345	87.966	F_3489	85.752	F_23579	86.384	F_234579	86.553		
F_346	88.380	F_3567	87.878	F_23589	85.711	F_234589	85.983		
F_347	88.319	F_3568	87.063	F_23678	87.396	F_234678	87.593		
F_348	87.484	F_3569	85.976	F_23679	86.295	F_234679	86.520		
F_349	86.180	F_3578	87.070	F_23689	85.576	F_234689	85.942		
F_356	87.864	F_3579	86.173	F_23789	85.657	F_234789	86.031		
F_357	87.762	F_3589	85.256	F_34567	88.027	F_235678	87.423		
F_358	86.941	F_3678	87.321	F_34568	87.246	F_235679	86.391		
F_359	85.969	F_3679	86.139	F_34569	86.241	F_235689	85.739		
F_367	88.394	F_3689	85.250	F_34578	87.233	F_235789	85.732		
F_368	87.124	F_3789	85.433	F_34579	86.404	F_236789	85.739		
F_369	85.827	F_12345	87.701	F_34589	85.725	F_345678	87.355		
F_378	87.199	F_12346	87.912	F_34678	87.559	F_345679	86.479		
F_379	86.153	F_12347	87.973	F_34679	86.526	F_345689	85.847		
F_389	85.107	F_12348	87.260	F_34689	85.745	F_345789	85.922		
F_1234	87.857	F_12349	86.316	F_34789	85.861	F_346789	85.881		
F_1235	87.654	F_12356	87.654	F_35678	87.233	F_356789	85.623		
F_1236	87.823	F_12357	87.579	F_35679	86.302	F_1234567	87.810		
F_1237	87.803	F_12358	87.117	F_35689	85.474	F_1234568	87.362		
F_1238	87.002	F_12359	86.098	F_35789	85.494	F_1234569	86.370		
F_1239	86.146	F_12367	87.891	F_36789	85.569	F_1234578	87.341		
F_1345	87.986	F_12368	87.239	F_123456	87.728	F_1234579	86.486		
F_1346	88.190	F_12369	86.180	F_123457	87.762	F_1234589	85.908		
F_1347	88.211	F_12378	87.267	F_123458	87.321	F_1234678	87.362		
F_1348	87.497	F_12379	86.228	F_123459	86.418	F_1234679	86.506		
F_1349	86.418	F_12389	85.460	F_123467	87.980	F_1234689	85.942		
F_1356	88.000	F_13456	88.068	F_123468	87.314	F_1234789	85.949		
F_1357	87.946	F_13457	87.966	F_123469	86.370	F_1235678	87.294		
F_1358	87.171	F_13458	87.280	F_123478	87.328	F_1235679	86.370		
F_1359	86.166	F_13459	86.452	F_123479	86.438	F_1235689	85.698		

Table 18. The complete results of concatenating the learned CNN features in the task CAL. This table shows that concatenating the CNN features learned from other tasks can improve the performance of CAL and that concatenating all the features does not achieve the best performance.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	48.845	F_1467	55.218	F_13467	55.035	F_123489	53.370	F_1245789	53.276
F_2	48.320	F_1468	53.930	F_13468	54.286	F_124567	54.718	F_1246789	53.048
F_3	50.411	F_1469	53.256	F_13469	53.568	F_124568	54.078	F_1345678	54.599
F_4	54.440	F_1478	54.301	F_13478	54.599	F_124569	53.563	F_1345679	53.806
F_5	50.629	F_1479	53.300	F_13479	53.538	F_124578	54.217	F_1345689	53.370
F_6	46.824	F_1489	52.572	F_13489	53.092	F_124579	53.721	F_1345789	53.503
F_7	48.697	F_2345	54.554	F_14567	55.382	F_124589	53.102	F_1346789	53.365
F_8	40.986	F_2346	54.539	F_14568	54.480	F_124678	54.058	F_1456789	53.300
F_9	43.925	F_2347	54.836	F_14569	53.399	F_124679	53.583	F_2345678	54.296
F_14	55.124	F_2348	53.954	F_14578	54.604	F_124689	52.810	F_2345679	53.736
F_24	54.574	F_2349	53.593	F_14579	53.523	F_124789	53.082	F_2345689	53.389
F_34	54.747	F_2456	54.564	F_14589	53.018	F_134567	55.208	F_2345789	53.290
F_45	55.149	F_2457	54.827	F_14678	54.306	F_134568	54.589	F_2346789	53.285
F_46	54.990	F_2458	54.039	F_14679	53.370	F_134569	53.850	F_2456789	53.385
F_47	55.064	F_2459	53.513	F_14689	52.572	F_134578	54.688	F_3456789	53.385
F_48	53.271	F_2467	54.936	F_14789	52.681	F_134579	53.717	F_12345678	54.267
F_49	52.359	F_2468	53.741	F_23456	54.693	F_134589	53.236	F_12345679	53.890
F_124	54.628	F_2469	53.256	F_23457	54.703	F_134678	54.633	F_12345689	53.444
F_134	55.000	F_2478	53.979	F_23458	54.237	F_134679	53.503	F_12345789	53.449
F_145	55.193	F_2479	53.385	F_23459	53.840	F_134689	53.107	F_12346789	53.449
F_146	55.253	F_2489	52.666	F_23467	54.822	F_134789	53.325	F_12456789	53.201
F_147	55.238	F_3456	55.040	F_23468	54.073	F_145678	54.569	F_13456789	53.394
F_148	53.761	F_3457	55.164	F_23469	53.598	F_145679	53.622	F_23456789	53.390
F_149	53.102	F_3458	54.425	F_23478	54.108	F_145689	53.231	F_123456789	53.489
F_234	54.554	F_3459	53.687	F_23479	53.707	F_145789	53.276		
F_245	54.618	F_3467	55.094	F_23489	53.132	F_146789	52.844		
F_246	54.792	F_3468	54.257	F_24567	54.807	F_234567	54.673		
F_247	54.876	F_3469	53.355	F_24568	54.143	F_234568	54.301		
F_248	53.583	F_3478	54.450	F_24569	53.548	F_234569	53.801		
F_249	53.142	F_3479	53.409	F_24578	54.252	F_234578	54.361		
F_345	55.005	F_3489	52.696	F_24579	53.548	F_234579	53.776		
F_346	55.015	F_4567	55.188	F_24589	52.973	F_234589	53.335		
F_347	55.084	F_4568	54.187	F_24678	54.004	F_234678	54.222		
F_348	54.163	F_4569	53.236	F_24679	53.498	F_234679	53.746		
F_349	53.300	F_4578	54.366	F_24689	52.716	F_234689	53.181		
F_456	55.228	F_4579	53.380	F_24789	52.978	F_234789	53.221		
F_457	55.277	F_4589	52.730	F_34567	55.178	F_245678	54.311		
F_458	54.158	F_4678	53.999	F_34568	54.425	F_245679	53.523		
F_459	53.176	F_4679	52.874	F_34569	53.741	F_245689	53.028		
F_467	55.213	F_4689	52.185	F_34578	54.544	F_245789	53.246		
F_468	53.513	F_4789	52.483	F_34579	53.746	F_246789	52.968		
F_469	52.612	F_12345	54.723	F_34589	53.256	F_345678	54.613		
F_478	53.959	F_12346	54.584	F_34678	54.440	F_345679	53.682		
F_479	52.780	F_12347	54.762	F_34679	53.474	F_345689	53.226		
F_489	51.947	F_12348	54.128	F_34689	52.755	F_345789	53.385		
F_1234	54.609	F_12349	53.707	F_34789	52.948	F_346789	53.127		
F_1245	54.564	F_12456	54.589	F_45678	54.485	F_456789	53.107		
F_1246	54.876	F_12457	54.861	F_45679	53.419	F_1234567	54.807		
F_1247	54.797	F_12458	54.004	F_45689	52.884	F_1234568	54.331		
F_1248	53.835	F_12459	53.608	F_45789	52.998	F_1234569	53.751		
F_1249	53.360	F_12467	54.886	F_46789	52.562	F_1234578	54.346		
F_1345	55.218	F_12468	53.885	F_123456	54.752	F_1234579	53.821		
F_1346	55.099	F_12469	53.503	F_123457	54.846	F_1234589	53.454		
F_1347	55.154	F_12478	54.058	F_123458	54.316	F_1234678	54.227		
F_1348	54.291	F_12479	53.523	F_123459	53.831	F_1234679	53.751		
F_1349	53.627	F_12489	52.775	F_123467	54.757	F_1234689	53.320		
F_1456	55.282	F_13456	55.193	F_123468	54.232	F_1234789	53.370		
F_1457	55.322	F_13457	55.164	F_123469	53.712	F_1245678	54.207		
F_1458	54.435	F_13458	54.564	F_123478	54.167	F_1245679	53.647		
F_1459	53.508	F_13459	53.786	F_123479	53.736	F_1245689	53.142		

Table 19. The complete results of concatenating the learned CNN features in the task ARC. This table shows that concatenating the CNN features learned from other tasks can improve the performance of ARC and that concatenating all the features does not achieve the best performance.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	31.138	F_1567	37.096	F_13567	37.467	F_123589	37.588	F_1245789	37.962
F_2	31.139	F_1568	38.461	F_13568	36.849	F_124567	37.094	F_1256789	36.223
F_3	31.886	F_1569	35.731	F_13569	36.848	F_124568	37.590	F_1345678	38.334
F_4	33.868	F_1578	38.089	F_13578	36.726	F_124569	36.970	F_1345679	36.473
F_5	36.228	F_1579	35.233	F_13579	35.730	F_124578	38.086	F_1345689	37.962
F_6	33.372	F_1589	36.723	F_13589	37.093	F_124579	36.723	F_1345789	38.210
F_7	33.748	F_2345	36.723	F_14567	38.709	F_124589	37.216	F_1356789	37.960
F_8	27.170	F_2356	37.468	F_14568	39.080	F_125678	36.472	F_1456789	38.209
F_9	30.029	F_2357	37.220	F_14569	36.845	F_125679	35.109	F_2345678	38.460
F_15	37.095	F_2358	37.470	F_14578	38.584	F_125689	35.604	F_2345679	36.473
F_25	36.600	F_2359	35.978	F_14579	36.846	F_125789	36.099	F_2345689	37.468
F_35	36.352	F_2456	37.714	F_14589	38.211	F_134567	38.337	F_2345789	37.839
F_45	35.855	F_2457	37.591	F_15678	37.965	F_134568	37.716	F_2356789	36.847
F_56	36.601	F_2458	37.095	F_15679	35.110	F_134569	37.467	F_2456789	37.216
F_57	36.105	F_2459	37.220	F_15689	36.972	F_134578	38.334	F_3456789	38.706
F_58	36.479	F_2567	36.102	F_15789	36.598	F_134579	36.349	F_12345678	37.838
F_59	35.361	F_2568	36.474	F_23456	36.351	F_134589	37.838	F_12345679	35.852
F_125	37.219	F_2569	35.608	F_23457	36.972	F_135678	36.850	F_12345689	37.841
F_135	37.716	F_2578	36.476	F_23458	38.213	F_135679	36.349	F_12345789	37.467
F_145	36.845	F_2579	35.855	F_23459	36.474	F_135689	37.591	F_12356789	37.463
F_156	37.220	F_2589	35.976	F_23567	36.723	F_135789	37.465	F_12456789	37.464
F_157	36.475	F_3456	36.723	F_23568	37.594	F_145678	38.210	F_13456789	38.087
F_158	38.090	F_3457	36.972	F_23569	36.723	F_145679	36.846	F_23456789	37.716
F_159	35.730	F_3458	37.841	F_23578	37.470	F_145689	38.087	F_123456789	36.971
F_235	37.467	F_3459	36.477	F_23579	35.483	F_145789	38.706		
F_245	36.845	F_3567	36.104	F_23589	37.468	F_156789	36.474		
F_256	36.599	F_3568	37.100	F_24567	37.715	F_234567	36.599		
F_257	36.351	F_3569	36.475	F_24568	37.343	F_234568	38.213		
F_258	35.731	F_3578	37.471	F_24569	36.847	F_234569	36.226		
F_259	35.607	F_3579	36.351	F_24578	37.220	F_234578	38.337		
F_345	36.972	F_3589	37.094	F_24579	37.343	F_234579	36.225		
F_356	36.724	F_4567	37.470	F_24589	37.963	F_234589	38.088		
F_357	36.723	F_4568	38.212	F_25678	36.725	F_235678	37.594		
F_358	37.347	F_4569	36.970	F_25679	35.855	F_235679	35.855		
F_359	36.972	F_4578	37.842	F_25689	36.101	F_235689	36.971		
F_456	37.220	F_4579	36.845	F_25789	36.102	F_235789	37.592		
F_457	35.857	F_4589	38.088	F_34567	37.220	F_245678	36.970		
F_458	39.082	F_5678	37.345	F_34568	37.842	F_245679	36.722		
F_459	37.591	F_5679	35.855	F_34569	36.599	F_245689	37.466		
F_567	36.354	F_5689	36.599	F_34578	38.213	F_245789	37.838		
F_568	37.223	F_5789	36.598	F_34579	36.599	F_256789	36.101		
F_569	35.733	F_12345	37.220	F_34589	38.334	F_345678	37.592		
F_578	37.098	F_12356	37.095	F_35678	37.223	F_345679	36.598		
F_579	34.861	F_12357	37.095	F_35679	36.351	F_345689	38.210		
F_589	36.228	F_12358	37.220	F_35689	37.217	F_345789	38.334		
F_1235	36.599	F_12359	35.605	F_35789	37.216	F_356789	37.341		
F_1245	37.963	F_12456	38.087	F_45678	37.592	F_456789	38.459		
F_1256	37.715	F_12457	37.218	F_45679	36.721	F_1234567	36.970		
F_1257	36.100	F_12458	37.963	F_45689	37.838	F_1234568	38.460		
F_1258	36.846	F_12459	37.095	F_45789	38.336	F_1234569	35.854		
F_1259	34.489	F_12567	36.598	F_56789	36.972	F_1234578	38.460		
F_1345	37.841	F_12568	37.095	F_123456	37.095	F_1234579	36.102		
F_1356	38.336	F_12569	34.986	F_123457	37.095	F_1234589	37.468		
F_1357	37.589	F_12578	36.845	F_123458	38.335	F_1235678	37.591		
F_1358	37.098	F_12579	35.234	F_123459	35.978	F_1235679	35.730		
F_1359	36.600	F_12589	36.349	F_123567	37.343	F_1235689	37.464		
F_1456	37.963	F_13456	38.213	F_123568	37.468	F_1235789	37.464		
F_1457	37.468	F_13457	37.964	F_123569	35.481	F_1245678	38.087		
F_1458	39.080	F_13458	38.089	F_123578	37.466	F_1245679	36.351		
F_1459	36.970	F_13459	37.220	F_123579	35.234	F_1245689	37.463		

Table 20. The complete results of concatenating the learned CNN features in the task EMO. This table shows that concatenating the CNN features learned from other tasks can improve the performance of EMO and that concatenating all the features does not achieve the best performance.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	63.181	F_1567	69.669	F_13567	69.528	F_123689	69.283	F_1246789	69.719
F_2	62.915	F_1568	69.834	F_13568	69.689	F_124567	69.503	F_1256789	69.508
F_3	63.297	F_1569	69.574	F_13569	69.192	F_124568	69.860	F_1345678	69.674
F_4	63.277	F_1678	69.809	F_13678	69.689	F_124569	69.443	F_1345679	69.584
F_5	63.392	F_1679	69.679	F_13679	69.513	F_124678	69.764	F_1345689	69.574
F_6	69.423	F_1689	69.980	F_13689	69.533	F_124679	69.523	F_1346789	69.634
F_7	63.006	F_2346	69.518	F_14567	69.629	F_124689	69.488	F_1356789	69.694
F_8	61.004	F_2356	69.468	F_14568	69.719	F_125678	69.729	F_1456789	69.759
F_9	61.666	F_2367	69.508	F_14569	69.513	F_125679	69.192	F_2345678	69.769
F_16	69.498	F_2368	69.624	F_14678	69.975	F_125689	69.368	F_2345679	69.333
F_26	69.528	F_2369	69.348	F_14679	69.533	F_126789	69.513	F_2345689	69.353
F_36	69.323	F_2456	69.483	F_14689	69.704	F_134567	69.609	F_2346789	69.237
F_46	69.503	F_2467	69.559	F_15678	69.809	F_134568	69.649	F_2356789	69.533
F_56	69.473	F_2468	69.644	F_15679	69.754	F_134569	69.473	F_2456789	69.538
F_67	69.729	F_2469	69.398	F_15689	69.839	F_134678	69.764	F_3456789	69.559
F_68	69.528	F_2567	69.609	F_16789	69.794	F_134679	69.569	F_12345678	69.719
F_69	69.594	F_2568	69.809	F_23456	69.729	F_134689	69.664	F_12345679	69.313
F_126	69.594	F_2569	69.463	F_23467	69.589	F_135678	69.719	F_12345689	69.438
F_136	69.288	F_2678	69.639	F_23468	69.553	F_135679	69.448	F_12346789	69.463
F_146	69.609	F_2679	69.594	F_23469	69.207	F_135689	69.634	F_12356789	69.463
F_156	69.553	F_2689	69.654	F_23567	69.574	F_136789	69.634	F_12456789	69.493
F_167	69.594	F_3456	69.453	F_23568	69.609	F_145678	69.850	F_13456789	69.569
F_168	69.880	F_3467	69.594	F_23569	69.157	F_145679	69.674	F_23456789	69.423
F_169	69.729	F_3468	69.744	F_23678	69.624	F_145689	69.799	F_123456789	69.458
F_236	69.468	F_3469	69.348	F_23679	69.398	F_146789	69.689		
F_246	69.358	F_3567	69.533	F_23689	69.087	F_156789	69.829		
F_256	69.423	F_3568	69.654	F_24567	69.498	F_234567	69.634		
F_267	69.684	F_3569	69.097	F_24568	69.845	F_234568	69.734		
F_268	69.594	F_3678	69.674	F_24569	69.338	F_234569	69.288		
F_269	69.408	F_3679	69.493	F_24678	69.689	F_234678	69.709		
F_346	69.428	F_3689	69.553	F_24679	69.498	F_234679	69.378		
F_356	69.363	F_4567	69.744	F_24689	69.483	F_234689	69.112		
F_367	69.564	F_4568	69.920	F_25678	69.614	F_235678	69.634		
F_368	69.569	F_4569	69.398	F_25679	69.418	F_235679	69.338		
F_369	69.313	F_4678	69.754	F_25689	69.513	F_235689	69.423		
F_456	69.488	F_4679	69.699	F_26789	69.599	F_236789	69.333		
F_467	69.814	F_4689	69.714	F_34567	69.403	F_245678	69.699		
F_468	69.744	F_5678	69.875	F_34568	69.709	F_245679	69.278		
F_469	69.443	F_5679	69.589	F_34569	69.383	F_245689	69.458		
F_567	69.589	F_5689	69.819	F_34678	69.915	F_246789	69.559		
F_568	69.624	F_6789	69.855	F_34679	69.579	F_256789	69.548		
F_569	69.413	F_12346	69.473	F_34689	69.458	F_345678	69.739		
F_678	69.935	F_12356	69.483	F_35678	69.734	F_345679	69.423		
F_679	69.523	F_12367	69.423	F_35679	69.413	F_345689	69.403		
F_689	69.764	F_12368	69.664	F_35689	69.448	F_346789	69.564		
F_1236	69.604	F_12369	69.473	F_36789	69.614	F_356789	69.458		
F_1246	69.498	F_12456	69.498	F_45678	69.779	F_456789	69.538		
F_1256	69.488	F_12467	69.589	F_45679	69.533	F_1234567	69.548		
F_1267	69.609	F_12468	69.513	F_45689	69.624	F_1234568	69.719		
F_1268	69.689	F_12469	69.478	F_46789	69.684	F_1234569	69.378		
F_1269	69.538	F_12567	69.458	F_56789	69.885	F_1234678	69.619		
F_1346	69.488	F_12568	69.719	F_123456	69.664	F_1234679	69.448		
F_1356	69.438	F_12569	69.358	F_123467	69.729	F_1234689	69.378		
F_1367	69.458	F_12678	69.724	F_123468	69.669	F_1235678	69.584		
F_1368	69.769	F_12679	69.594	F_123469	69.433	F_1235679	69.383		
F_1369	69.453	F_12689	69.649	F_123567	69.463	F_1235689	69.488		
F_1456	69.604	F_13456	69.523	F_123568	69.639	F_1236789	69.498		
F_1467	69.784	F_13467	69.513	F_123569	69.368	F_1245678	69.860		
F_1468	69.865	F_13468	69.759	F_123678	69.679	F_1245679	69.383		
F_1469	69.518	F_13469	69.669	F_123679	69.483	F_1245689	69.599		

Table 21. The complete results of concatenating the learned CNN features in the task AVA. This table shows that concatenating the CNN features learned from other tasks can improve the performance of AVA and that concatenating all the features does not achieve the best performance.

feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy	feature	accuracy
F_1	68.478	F_1567	76.087	F_13567	77.174	F_123789	74.565	F_1246789	74.348
F_2	66.739	F_1578	76.304	F_13578	76.522	F_124567	75.217	F_1256789	75.217
F_3	66.739	F_1579	74.565	F_13579	75.000	F_124578	75.652	F_1345678	76.087
F_4	71.739	F_1678	77.391	F_13678	76.522	F_124579	74.783	F_1345679	75.870
F_5	68.043	F_1679	75.652	F_13679	75.435	F_124678	76.087	F_1345789	75.217
F_6	69.565	F_1789	74.565	F_13789	74.783	F_124679	74.783	F_1346789	75.217
F_7	76.957	F_2347	76.957	F_14567	75.870	F_124789	74.783	F_1356789	75.870
F_8	60.870	F_2357	75.217	F_14578	76.957	F_125678	74.565	F_1456789	75.652
F_9	66.087	F_2367	74.130	F_14579	75.000	F_125679	73.478	F_2345678	75.435
F_17	76.957	F_2378	75.000	F_14678	76.957	F_125789	74.565	F_2345679	75.435
F_27	74.348	F_2379	73.261	F_14679	74.130	F_126789	74.783	F_2345789	74.565
F_37	76.522	F_2457	75.652	F_14789	75.000	F_134567	75.652	F_2346789	74.565
F_47	75.652	F_2467	76.739	F_15678	76.957	F_134578	76.957	F_2356789	75.217
F_57	75.435	F_2478	73.913	F_15679	75.217	F_134579	76.087	F_2456789	75.435
F_67	77.391	F_2479	74.783	F_15789	74.783	F_134678	75.870	F_3456789	75.000
F_78	76.957	F_2567	76.087	F_16789	74.783	F_134679	75.870	F_12345678	75.435
F_79	74.348	F_2578	74.130	F_23457	75.652	F_134789	75.000	F_12345679	74.348
F_127	74.348	F_2579	74.565	F_23467	75.652	F_135678	76.739	F_12345789	74.565
F_137	76.087	F_2678	74.565	F_23478	75.000	F_135679	75.870	F_12346789	74.783
F_147	75.870	F_2679	74.130	F_23479	75.217	F_135789	75.652	F_12356789	74.783
F_157	76.304	F_2789	74.565	F_23567	74.783	F_136789	75.000	F_12456789	75.000
F_167	76.304	F_3457	75.435	F_23578	75.435	F_145678	77.174	F_13456789	75.435
F_178	76.522	F_3467	76.522	F_23579	73.913	F_145679	75.000	F_23456789	74.565
F_179	74.783	F_3478	76.087	F_23678	75.217	F_145789	75.435	F_123456789	74.348
F_237	75.000	F_3479	76.522	F_23679	73.913	F_146789	74.348		
F_247	76.087	F_3567	76.522	F_23789	75.000	F_156789	74.348		
F_257	75.870	F_3578	75.000	F_24567	76.304	F_234567	75.652		
F_267	75.217	F_3579	75.217	F_24578	73.696	F_234578	75.217		
F_278	74.565	F_3678	75.652	F_24579	74.783	F_234579	75.435		
F_279	73.913	F_3679	74.565	F_24678	74.565	F_234678	75.870		
F_347	76.957	F_3789	74.565	F_24679	74.348	F_234679	74.348		
F_357	76.087	F_4567	76.087	F_24789	75.000	F_234789	74.348		
F_367	75.652	F_4578	75.652	F_25678	74.348	F_235678	74.565		
F_378	75.435	F_4579	75.652	F_25679	74.783	F_235679	74.348		
F_379	74.130	F_4678	75.435	F_25789	74.348	F_235789	75.217		
F_457	75.435	F_4679	74.348	F_26789	74.565	F_236789	75.000		
F_467	75.652	F_4789	75.000	F_34567	75.217	F_245678	74.130		
F_478	76.522	F_5678	76.957	F_34578	76.739	F_245679	74.783		
F_479	74.348	F_5679	75.217	F_34579	75.652	F_245789	75.000		
F_567	76.087	F_5789	75.217	F_34678	75.870	F_246789	74.783		
F_578	77.391	F_6789	74.565	F_34679	76.522	F_256789	75.435		
F_579	74.783	F_12347	76.957	F_34789	75.000	F_345678	76.522		
F_678	76.087	F_12357	76.087	F_35678	75.435	F_345679	75.652		
F_679	74.130	F_12367	75.217	F_35679	75.217	F_345789	75.435		
F_789	74.348	F_12378	74.565	F_35789	75.652	F_346789	75.000		
F_1237	74.565	F_12379	73.261	F_36789	75.217	F_356789	76.087		
F_1247	75.652	F_12457	75.435	F_45678	75.652	F_456789	75.870		
F_1257	75.870	F_12467	75.652	F_45679	75.870	F_1234567	75.870		
F_1267	74.783	F_12478	75.870	F_45789	75.652	F_1234578	75.652		
F_1278	75.870	F_12479	74.565	F_46789	75.435	F_1234579	74.348		
F_1279	74.565	F_12567	75.870	F_56789	75.217	F_1234678	74.783		
F_1347	77.174	F_12578	75.217	F_123457	76.087	F_1234679	74.565		
F_1357	77.609	F_12579	73.044	F_123467	76.739	F_1234789	75.000		
F_1367	76.304	F_12678	75.000	F_123478	75.870	F_1235678	75.435		
F_1378	77.174	F_12679	74.348	F_123479	74.348	F_1235679	73.913		
F_1379	75.435	F_12789	75.217	F_123567	76.087	F_1235789	75.217		
F_1457	75.652	F_13457	76.304	F_123578	75.435	F_1236789	74.783		
F_1467	75.870	F_13467	76.957	F_123579	73.261	F_1245678	75.435		
F_1478	77.174	F_13478	76.957	F_123678	74.783	F_1245679	75.435		
F_1479	74.348	F_13479	76.087	F_123679	73.261	F_1245789	75.435		

Table 22. The complete results of concatenating the learned CNN features in the task FAS. This table shows that concatenating the CNN features learned from other tasks can improve the performance of FAS and that concatenating all the features does not achieve the best performance.

feature	ρ	feature	ρ	feature	ρ	feature	ρ	feature	ρ
F_1	0.454	F_1568	0.494	F_13568	0.493	F_123789	0.494	F_1246789	0.495
F_2	0.442	F_1578	0.493	F_13578	0.494	F_124568	0.505	F_1256789	0.500
F_3	0.464	F_1589	0.485	F_13589	0.490	F_124578	0.505	F_1345678	0.501
F_4	0.434	F_1678	0.487	F_13678	0.490	F_124589	0.497	F_1345689	0.496
F_5	0.459	F_1689	0.470	F_13689	0.483	F_124678	0.501	F_1345789	0.497
F_6	0.445	F_1789	0.474	F_13789	0.486	F_124689	0.491	F_1346789	0.493
F_7	0.450	F_2348	0.495	F_14568	0.499	F_124789	0.494	F_1356789	0.496
F_8	0.398	F_2358	0.495	F_14578	0.499	F_125678	0.503	F_1456789	0.497
F_9	0.346	F_2368	0.489	F_14589	0.490	F_125689	0.497	F_2345678	0.505
F_18	0.459	F_2378	0.492	F_14678	0.495	F_125789	0.498	F_2345689	0.500
F_28	0.466	F_2389	0.485	F_14689	0.480	F_126789	0.490	F_2345789	0.501
F_38	0.468	F_2458	0.498	F_14789	0.484	F_134568	0.499	F_2346789	0.498
F_48	0.463	F_2468	0.494	F_15678	0.497	F_134578	0.499	F_2356789	0.500
F_58	0.470	F_2478	0.495	F_15689	0.490	F_134589	0.494	F_2456789	0.500
F_68	0.465	F_2489	0.482	F_15789	0.491	F_134678	0.496	F_3456789	0.497
F_78	0.461	F_2568	0.497	F_16789	0.480	F_134689	0.489	F_12345678	0.506
F_89	0.417	F_2578	0.496	F_23458	0.501	F_134789	0.491	F_12345689	0.502
F_128	0.482	F_2589	0.488	F_23468	0.498	F_135678	0.496	F_12345789	0.503
F_138	0.479	F_2678	0.490	F_23478	0.499	F_135689	0.493	F_12346789	0.499
F_148	0.482	F_2689	0.476	F_23489	0.492	F_135789	0.494	F_12356789	0.502
F_158	0.488	F_2789	0.479	F_23568	0.498	F_136789	0.489	F_12456789	0.502
F_168	0.478	F_3458	0.492	F_23578	0.498	F_145678	0.502	F_13456789	0.499
F_178	0.479	F_3468	0.489	F_23589	0.494	F_145689	0.493	F_23456789	0.502
F_189	0.457	F_3478	0.491	F_23678	0.494	F_145789	0.495	F_123456789	0.504
F_238	0.485	F_3489	0.483	F_23689	0.489	F_146789	0.487		
F_248	0.488	F_3568	0.487	F_23789	0.491	F_156789	0.494		
F_258	0.491	F_3578	0.488	F_24568	0.502	F_234568	0.503		
F_268	0.481	F_3589	0.485	F_24578	0.502	F_234578	0.503		
F_278	0.483	F_3678	0.486	F_24589	0.494	F_234589	0.498		
F_289	0.466	F_3689	0.480	F_24678	0.499	F_234678	0.501		
F_348	0.484	F_3789	0.484	F_24689	0.487	F_234689	0.494		
F_358	0.481	F_4568	0.494	F_24789	0.490	F_234789	0.496		
F_368	0.479	F_4578	0.494	F_25678	0.500	F_235678	0.500		
F_378	0.481	F_4589	0.483	F_25689	0.492	F_235689	0.497		
F_389	0.474	F_4678	0.490	F_25789	0.494	F_235789	0.498		
F_458	0.487	F_4689	0.473	F_26789	0.485	F_236789	0.493		
F_468	0.480	F_4789	0.478	F_34568	0.496	F_245678	0.505		
F_478	0.482	F_5678	0.490	F_34578	0.496	F_245689	0.496		
F_489	0.461	F_5689	0.484	F_34589	0.491	F_245789	0.498		
F_568	0.486	F_5789	0.484	F_34678	0.494	F_246789	0.493		
F_578	0.482	F_6789	0.473	F_34689	0.487	F_256789	0.497		
F_589	0.474	F_12348	0.498	F_34789	0.489	F_345678	0.498		
F_678	0.480	F_12358	0.499	F_35678	0.492	F_345689	0.494		
F_689	0.455	F_12368	0.494	F_35689	0.489	F_345789	0.495		
F_789	0.460	F_12378	0.495	F_35789	0.491	F_346789	0.492		
F_1238	0.491	F_12389	0.490	F_36789	0.487	F_356789	0.493		
F_1248	0.495	F_12458	0.503	F_45678	0.498	F_456789	0.494		
F_1258	0.498	F_12468	0.498	F_45689	0.489	F_1234568	0.505		
F_1268	0.489	F_12478	0.499	F_45789	0.490	F_1234578	0.505		
F_1278	0.491	F_12489	0.488	F_46789	0.484	F_1234589	0.500		
F_1289	0.478	F_12568	0.501	F_56789	0.489	F_1234678	0.502		
F_1348	0.489	F_12578	0.501	F_123458	0.503	F_1234689	0.496		
F_1358	0.490	F_12589	0.494	F_123468	0.500	F_1234789	0.498		
F_1368	0.484	F_12678	0.495	F_123478	0.501	F_1235678	0.503		
F_1378	0.487	F_12689	0.484	F_123489	0.495	F_1235689	0.499		
F_1389	0.479	F_12789	0.487	F_123568	0.501	F_1235789	0.501		
F_1458	0.496	F_13458	0.496	F_123578	0.502	F_1236789	0.496		
F_1468	0.489	F_13468	0.492	F_123589	0.498	F_1245678	0.507		
F_1478	0.491	F_13478	0.494	F_123678	0.497	F_1245689	0.499		
F_1489	0.474	F_13489	0.486	F_123689	0.492	F_1245789	0.501		

Table 23. The complete results of concatenating the learned CNN features in the task MEM. ρ is the Spearman rank correlation between the prediction and the ground truth. This table shows that concatenating the CNN features learned from other tasks can improve the performance of MEM and that concatenating all the features does not achieve the best performance.

feature	ρ	feature	ρ	feature	ρ	feature	ρ	feature	ρ
F_1	0.520	F_1569	0.618	F_13569	0.619	F_123789	0.627	F_1246789	0.628
F_2	0.508	F_1579	0.621	F_13579	0.624	F_124569	0.623	F_1256789	0.626
F_3	0.493	F_1589	0.616	F_13589	0.619	F_124579	0.623	F_1345679	0.625
F_4	0.487	F_1679	0.621	F_13679	0.628	F_124589	0.625	F_1345689	0.625
F_5	0.497	F_1689	0.615	F_13689	0.623	F_124679	0.625	F_1345789	0.626
F_6	0.575	F_1789	0.619	F_13789	0.626	F_124689	0.627	F_1346789	0.630
F_7	0.542	F_2349	0.623	F_14569	0.621	F_124789	0.627	F_1356789	0.626
F_8	0.560	F_2359	0.619	F_14579	0.623	F_125679	0.624	F_1456789	0.626
F_9	0.573	F_2369	0.621	F_14589	0.624	F_125689	0.623	F_2345679	0.626
F_19	0.598	F_2379	0.624	F_14679	0.625	F_125789	0.625	F_2345689	0.626
F_29	0.599	F_2389	0.619	F_14689	0.624	F_126789	0.626	F_2345789	0.626
F_39	0.598	F_2459	0.616	F_14789	0.627	F_134569	0.622	F_2346789	0.629
F_49	0.598	F_2469	0.617	F_15679	0.624	F_134579	0.624	F_2356789	0.625
F_59	0.599	F_2479	0.619	F_15689	0.619	F_134589	0.624	F_2456789	0.625
F_69	0.610	F_2489	0.620	F_15789	0.623	F_134679	0.628	F_3456789	0.624
F_79	0.603	F_2569	0.615	F_16789	0.623	F_134689	0.628	F_12345679	0.629
F_89	0.585	F_2579	0.618	F_23459	0.621	F_134789	0.628	F_12345689	0.628
F_129	0.612	F_2589	0.613	F_23469	0.624	F_135679	0.625	F_12345789	0.629
F_139	0.615	F_2679	0.618	F_23479	0.625	F_135689	0.622	F_12346789	0.630
F_149	0.616	F_2689	0.615	F_23489	0.626	F_135789	0.624	F_12356789	0.628
F_159	0.613	F_2789	0.618	F_23569	0.621	F_136789	0.628	F_12456789	0.628
F_169	0.611	F_3459	0.614	F_23579	0.624	F_145679	0.624	F_13456789	0.627
F_179	0.618	F_3469	0.618	F_23589	0.619	F_145689	0.625	F_23456789	0.627
F_189	0.607	F_3479	0.619	F_23679	0.626	F_145789	0.625	F_123456789	0.629
F_239	0.617	F_3489	0.617	F_23689	0.623	F_146789	0.628		
F_249	0.613	F_3569	0.610	F_23789	0.624	F_156789	0.625		
F_259	0.611	F_3579	0.616	F_24569	0.618	F_234569	0.623		
F_269	0.611	F_3589	0.609	F_24579	0.620	F_234579	0.625		
F_279	0.613	F_3679	0.623	F_24589	0.622	F_234589	0.625		
F_289	0.610	F_3689	0.615	F_24679	0.621	F_234679	0.627		
F_349	0.612	F_3789	0.618	F_24689	0.623	F_234689	0.628		
F_359	0.604	F_4569	0.617	F_24789	0.623	F_234789	0.627		
F_369	0.610	F_4579	0.618	F_25679	0.621	F_235679	0.625		
F_379	0.619	F_4589	0.619	F_25689	0.616	F_235689	0.622		
F_389	0.608	F_4679	0.617	F_25789	0.619	F_235789	0.624		
F_459	0.613	F_4689	0.617	F_26789	0.622	F_236789	0.626		
F_469	0.609	F_4789	0.619	F_34569	0.617	F_245679	0.622		
F_479	0.613	F_5679	0.620	F_34579	0.618	F_245689	0.623		
F_489	0.614	F_5689	0.614	F_34589	0.619	F_245789	0.624		
F_569	0.609	F_5789	0.617	F_34679	0.622	F_246789	0.625		
F_579	0.615	F_6789	0.621	F_34689	0.621	F_256789	0.621		
F_589	0.608	F_12349	0.625	F_34789	0.622	F_345679	0.621		
F_679	0.619	F_12359	0.622	F_35679	0.618	F_345689	0.620		
F_689	0.612	F_12369	0.623	F_35689	0.613	F_345789	0.621		
F_789	0.612	F_12379	0.627	F_35789	0.617	F_346789	0.624		
F_1239	0.622	F_12389	0.624	F_36789	0.622	F_356789	0.619		
F_1249	0.619	F_12459	0.621	F_45679	0.621	F_456789	0.623		
F_1259	0.617	F_12469	0.622	F_45689	0.621	F_1234569	0.626		
F_1269	0.618	F_12479	0.623	F_45789	0.621	F_1234579	0.628		
F_1279	0.620	F_12489	0.625	F_46789	0.621	F_1234589	0.627		
F_1289	0.617	F_12569	0.621	F_56789	0.621	F_1234679	0.629		
F_1349	0.622	F_12579	0.623	F_123459	0.625	F_1234689	0.630		
F_1359	0.615	F_12589	0.620	F_123469	0.626	F_1234789	0.630		
F_1369	0.620	F_12679	0.624	F_123479	0.628	F_1235679	0.627		
F_1379	0.626	F_12689	0.622	F_123489	0.628	F_1235689	0.626		
F_1389	0.619	F_12789	0.623	F_123569	0.624	F_1235789	0.627		
F_1459	0.619	F_13459	0.620	F_123579	0.627	F_1236789	0.630		
F_1469	0.620	F_13469	0.624	F_123589	0.624	F_1245679	0.625		
F_1479	0.623	F_13479	0.626	F_123679	0.628	F_1245689	0.626		
F_1489	0.623	F_13489	0.626	F_123689	0.627	F_1245789	0.627		

Table 24. The complete results of concatenating the learned CNN features in the task INT. ρ is the Spearman rank correlation between the prediction and the ground truth. This table shows that concatenating the CNN features learned from other tasks can improve the performance of INT and that concatenating all the features does not achieve the best performance.