

Instruction pipeline and requirements of HW2

The following is a simple pipeline:

1. Compute SIFT for all the images. We suggest to save the features on disks after each iteration to avoid memory issue.
2. For all the features, apply with the K-Means algorithm. In this step, pay attention to the number of clusters and the number of max_iter.
3. Use the BoW model to quantize features and represent by frequencies.
4. Implement the SPM method.
5. Running SVM for image classification. To prevent 'out of memory' issue, you can load mini-batch of training samples from disks in each iteration for training. And you just need to implement with linear kernel, using histogram intersection kernel is a bonus.

Your report should include your results, how to run the code, problems you found, your solution and your new findings, etc.

Also, please use the CVPR template(which we have provided in github classrom), other formats are not acceptable.

Submission

Here are the things you need to submit:

1. A PDF-formatted report which describes what you have done ->Gradescope.
2. Homework code for hw2 -> github classroom.

For the first part, you should add CS172 course in your Gradescope. The course entry code is X355ND. Later we will add hw2 report submission channel so that you can submit your report. For the second part, you should accept this link <https://classroom.github.com/a/j2-OHHDR> to get the repo, which contains a README.md and a report template. You SHOULD fill your name and your student ID into README.md, otherwise we can't identify to whom this repo belongs and may give you 0 score for this homework. Please fill in your name and ID as soon as you clone the repo so that things mentioned above won't happen. Besides, you should include all your code into a zip file, naming it as hw2_studentID.zip(for example: hw2_2017123456.zip), so that we can keep a copy of your homework.

DDL: Oct 30, 23:59, 2021

hw2

Updated 1 month ago by Ruoyu Wang and Chenyu Wang

followup discussions *for lingering questions and comments*

☒ Resolved ☐ Unresolved



Anonymous Mouse 1 month ago

Do we need to implement all the steps? Can we use some libraries in some of the steps?

helpful! | 0



Chenyu Wang 1 month ago You can use some libraries to extract sift, implement K-means and SVM algorithm.

good comment | 0



Anonymous Gear 1 month ago Are all the funcs in libs allowed? Or are there any limitations on them?

helpful! | 0



Chenyu Wang 1 month ago You can use libs only to implement the steps mentioned above, as for which libs you want to use, there does not have limits. I don't know if I have a clear answer, if you have any questions, you're welcome to ask.

good comment | 0



Anonymous Gear 1 month ago For example: cv2.kmeans()

helpful! | 0



Chenyu Wang 1 month ago yeah, you can.

good comment | 0

☒ Resolved ☐ Unresolved



连奕航 1 month ago

For `Caltech256/138.mattress/138_0145.jpg` in the data set, it is strange that I cannot extract any SIFT feature from it. How can I deal with it? Can I remove it from the training set?

Thanks.

helpful! | 0



Chenyu Wang 1 month ago Can you offer some detailed information?

good comment | 0



连奕航 1 month ago

This image is shown below, which seems to have few obvious key features:



I found that, when I read this image with `cv2.imread` directly, SIFT can extract 2 descriptors from it; however, if convert it into an RGB image using `cv2.cvtColor(image, cv2.COLOR_BGR2RGB)`, `SIFT.detectAndCompute()` will return empty list of keypoints or descriptors for this image. How should I pre-process data to avoid this problem?

helpful! | 0



Ruoyu Wang 1 month ago It's really a strange image. But I can always find 2 keypoints by `SIFT.detectAndCompute()` whether I use GRAY, RGB or BGR. Maybe you should check the code or the version of opencv, or just ignore this image if you are really upset because of it.

good comment | 0



连奕航 1 month ago

Thanks! BTW, should I resize each image from the data set to the same size before building BOW or SPM?

helpful! | 0



Ruoyu Wang 1 month ago You can try and choose the way you like.

good comment | 0

☒ Resolved ☐ Unresolved



Anonymous Mouse 1 month ago

How many classes should we train?

helpful! | 0



Chenyu Wang 1 month ago All. But if your computer cannot support that, you can choose some of the classes. Though the scoring standard will be related to the number of the classes you used, it only account for a small proportion, so don't worry about that.

good comment | 0



Anonymous Helix 1 month ago

You mean we need to implement a multi-class SVM classifier?

helpful! | 0



Chenyu Wang 1 month ago yes

good comment | 0

☒ Resolved ☐ Unresolved



Anonymous Calc 1 month ago

Must we use LibSVM or we can use any SVM library we want?

helpful! | 0



Chenyu Wang 1 month ago That's OK.

good comment | 0

☒ Resolved ☐ Unresolved



Anonymous Poet 1 month ago

Must the number of clusters in k-means be equal to the number of classes? Or, how can we choose the parameter *num_clusters*?

helpful! | 0



Ruoyu Wang 1 month ago You decide based on what you learned.

good comment | 0

☒ Resolved ☐ Unresolved



Anonymous Beaker 1 month ago

Should we train 2 model using BoW representation only and SPM representation, or 1 model using SPM?

helpful! | 0



Chenyu Wang 1 month ago 2 model, one with BoW + SVM, and the other with BoW + SPM + SVM.

good comment | 0

☒ Resolved ☐ Unresolved



吴沁蔓 1 month ago

I wonder which dataset should we use? CIFAR10 or other dataset?

helpful! | 0



Anonymous Poet 1 month ago caltech256

helpful! | 0

☒ Resolved ☐ Unresolved



吴宗祥 1 month ago

Should I split the images in each classes into three parts, train, validation and test? If so, will the bag of features be generated just from the train set?

BTW, I wonder if I can use vlfeat in matlab, since it is mentioned in lec12? For example, vl_kmeans, vl_svmtrain and vl_kdtree.

helpful! | 0



Chenyu Wang 1 month ago You only need to split into train part and test part.

Yeah, you can feel free to use that.

good comment | 0



吴沁蔓 1 month ago So should the BOW be generated just from the train set?

helpful! | 1



Chenyu Wang 1 month ago @18_f9

good comment | 0



吴宗祥 1 month ago vl_dsift is TRASH!

helpful! | 0

☒ Resolved ☐ Unresolved



Anonymous Comp 1 month ago Is that means that we need to represent all the pictures with BOW, no matter test or train set, and use images' BOW representation as the input of SVM instead of images' pixels or images features?
helpful! | 0



Chenyu Wang 1 month ago Correct.

good comment | 0

☒ Resolved ☐ Unresolved



Anonymous Scale 1 month ago Is there any reference accuracy can be offered to us?
For example, what range should the accuracy be in when N classes, K clusters, n1 train images per class, bow + linear svm?

It will help me to tell whether my code is broken cause my results are some low accuracies and I've changed the parameters for several days. 😞

helpful! | 1



Anonymous Poet 1 month ago May we know how much will the accuracy affect our score? Thanks.

helpful! | 0



Chenyu Wang 1 month ago We mainly focus on your understanding of the whole process, so neither the accuracy nor the number of training data used would affect the score too much. Detailed grading standard has not decided.

good comment | 0

☒ Resolved ☐ Unresolved



Anonymous Scale 1 month ago Is there any reference accuracy can be offered to us?
For example, what range should the accuracy be in when N classes, K clusters, n1 train images per class, bow + linear svm?

It will help me to tell whether my code is broken cause my results are some low accuracies and I've changed the parameters for several days. 😞

helpful! | 0

☒ Resolved ☐ Unresolved



曾子尧 1 month ago

The ppt writes that "training number for each class: 15, 30, 45 60", does it means that we train the model with only these numbers of training data?

And how to set test set? Could we choose one third of training data to form test set(certainly those data is not the same of training data), like "5,19,15,20"?

Thank you for your time~

helpful! | 0



Chenyu Wang 1 month ago The first one, take 15 as an example, you should choose 15 samples for each class, and those samples you chosen make up the training set.

The second question, except those data used for training set, others remained are all test set.

good comment | 1

☒ Resolved ☐ Unresolved



吴宗祥 1 month ago

To classmates using matlab and vlfeat:

DON'T use `vl_sift` or `vl_dsift`!!!!

I used `vl_dsift` in the my code, and it demands images must be of class **SINGLE**. I've never thought about that our images will turn to **white images** after being changed from double to single!!!!!!!!!!!!

It means I've tested my code for two days, changed all the parameters for thousand of times but all my 128d features were [1,1,1,1,1,.....,1] and my accuracy was always ~10%!!!!

Hope this will help you. DON'T BE WRONG AT THE BEGINNING.

ps. `vl_kmeans`, `vl_kdtree` are safe

helpful! | 3

☒ Resolved ☐ Unresolved



曾子尧 1 month ago

Should we (or could we) discuss our method limitation in our report, just as CVPR 2022 newly requires this year?

Discussion of Limitations: Considering the limitations of an approach is an important part of good academic scholarship. While many high-quality CVPR papers thoroughly consider and discuss limitations, CVPR 2022 aims to emphasize the importance of an honest treatment of limitations. Authors are thus explicitly encouraged to discuss limitations, ideally by creating a separate "Limitations" (sub-)section in their papers.

Such discussion shall point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only held locally). Authors need to reflect on how these assumptions might be violated in practice and what the implications would be.

The discussion shall also reflect on the scope of their claims, e.g., if the authors only tested their approach on a few datasets or did a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated. The discussion shall reflect on the factors that influence the performance of the approach. For example, a recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting.

We understand that authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection. It is worth keeping in mind that a worse outcome might be if reviewers discover limitations that are not acknowledged in the paper. In general, we advise authors to use their best judgement and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations

helpful! | 0



Chenyu Wang 1 month ago It is not required.

good comment | 0