Automatic staging of placental maturity based on dense descriptor

Xinyao Li[a], Yuan Yao[b], Dong Ni[a], Siping Chen[a], Shengli Li[b], Baiying Lei[a,*] and Tianfu Wang[a,*]

[a]*Department of Biomedical Engineering, School of Medicine, Shenzhen University, National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Shenzhen, China*
[b]*Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare, Hospital of Nanfang Medical University, Shenzhen, China*

Abstract. Currently, placental maturity staging is mainly based on subjective observation of the physician. To address this issue, a new method is proposed for automatic staging of placental maturity based on B-mode ultrasound images. Due to small variations in the placental images, dense descriptor is utilized in place of the sparse descriptor to boost performance. Dense sampled DAISY descriptor is investigated for the demonstrated scale and translation invariant properties. Moreover, the extracted dense features are encoded by vector locally aggregated descriptor (VLAD) for performance boosting. The experimental results demonstrate an accuracy of 0.874, a sensitivity of 0.996 and a specificity of 0.874 for placental maturity staging. The experimental results also show that the dense features outperform the sparse features.

Keywords: Placental staging, daisy descriptor, dense sampling, vector locally aggregated descriptor

1. Introduction

Due to non-radiation, direct use and low cost, ultrasound imaging (US) has been broadly applied for prenatal diagnosis and staging. B-mode ultra-sound imaging is commonly used to assess placental abnormalities such as stillbirth, fetal death, small gestational age and other complications during pregnancy. Placental function assessment is a direct indication of fetal placental maturity. Currently, the assessment for placental function is based on the clinical judgment of the physician. Therefore, due to such subjective nature for diagnosis, discrepancies in assessment may arise depending on the clinical background of the physician. Moreover, since there is often a short supply of experienced physicians in certain countries, objective, automatic means of detection are needed. Furthermore,

objective diagnosis, together with the physician's subjective assessment can improve and validate the diagnosis.

There are a myriad of studies aimed at developing ultrasound based placental maturity assessments. For instance, Grammum [1] in 1979 first proposed to classify the chorionic plate, substance and basal plate of the placenta into four levels. Although this method achieved promising performance, it relied too much on the subjective interpretation of ultrasound images. In [2], Liu et al. purposed to automatically classify placental maturity based on SVM method. The proposed algorithm achieved a classification rate of 90% with quantitative features such as gray variance, distortions, and kurtosis. However, the classification results were not high enough for practical applications. To the best of our knowledge, there is still no practical application of placental maturity staging by automatic means.

To develop a new staging algorithm for placental maturity, points of interest should be detected first, followed by calculation of invariant feature descriptors based on these points. In the literature, numerous affine interest point available for covariant region detection. The commonly applied method includes Harris [3] and difference of Gaussian (DoG) [4]. Meanwhile, one of the most popular feature descriptor contains the commonly adopted [4] SIFT descriptor due to its invariance to a lot of modifications. In general, the derivatives and moment based descriptor achieves better result than the shape descriptor. For placental ultrasound images, there are a lot of modifications and changes such as complex illumination, exposure time and specular reflection by the imaging process, all of which makes placental ultrasound image staging challenging. The sparse method with the SIFT descriptor may not obtain good placental results. On the other hand, the dense sampling with the dense feature descriptor has demonstrated very good performance by its dense selection of discriminative features. Accordingly, the dense descriptor has been applied to object classifications [5]. In our method, the scale and translation invariant DAISY descriptor is utilized due to its demonstrated performance [6]. After the extraction of dense features, feature dimensions are often too high for the following classification task, which will cause classification burden due to a large number of redundant features. Therefore, feature learning and selection for dimension reduction are necessary for practical applications. Feature encoding is an effective way to reduce the feature dimension and boost the classification performance. The typical feature encoding methods are the bag of visual words (BoVW) [7], and aggregated codes of BoVW extensions such as vector of locally aggregated descriptor (VLAD) [8]. VLAD is proven to outperform the BoVW method by introducing higher statistics such as derivatives. Because of this, the VLAD feature encoding method is included in this paper for performance boosting. Furthermore, a histogram of occurrence is incorporated into the system to further enhance staging performance. Specifically, the final feature vector is formed by histogram obtained from the encoded feature vectors.

In this paper, a new placental staging method using the B-mode gray-scale ultrasound images are

Table1

Characteristics of 4 stages of placental maturity

Stage	Chorionic plate	Substance	Basal layer
0	Straight, smooth and chiseled	uniform	No echo
1	Slight undulating	unvenly distributed, scattered, point-like	No echo
2	In a serrated form, may extend into the substance of the placenta, but not the basal layer	Linearly echogenic, comma-like densities	Linear aligned, point-like echo
3	Jaggered, stretched into basal layer	Circular densities, halo with cast acoustic shadow	Large, confluent with basal layer,Acoustic shadow

proposed based on dense descriptor. The challenge of the placental image maturity staging is the lack of benchmark and uniform standard in clinical practices. The automatic staging of placental maturity is based on the gestational stages of the placental chorionic plate. Table 1 provides the 4-stages of placenta maturity [1,9] according to placenta variations, chorionic plate, placental substance and basal layer. The main aim of this paper is to design an automatic staging method for placental maturity evaluation using the staging benchmark features outlined in Table 1.

2. Methodology

2.1. System architecture

Figure 1 shows the system architecture of the proposed automatic staging of placental maturity. Dense sampling is applied to the preprocessed placental images, the DAISY features are extracted based on dense grid sampling. The extracted dense features are encoded by VLAD to reduce the feature dimension and save memory space. The densely sampled feature vectors are constructed in Figure 2. After extraction of patches, each feature is extracted by the local region invariant DAISY descriptor.
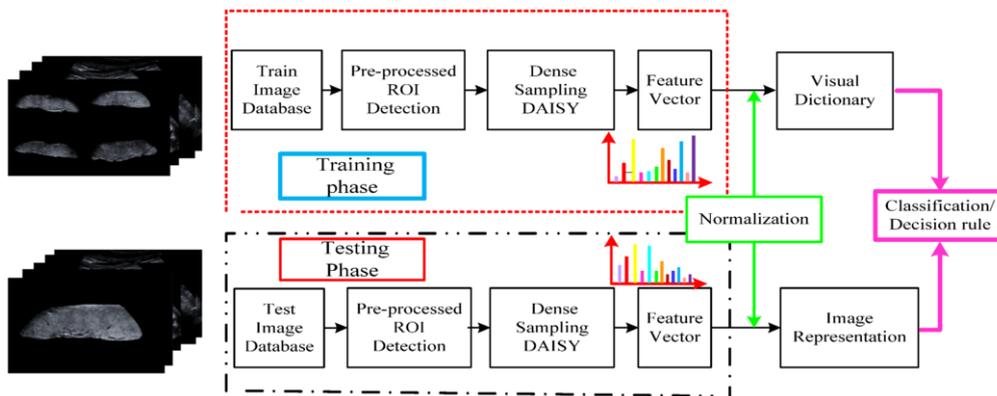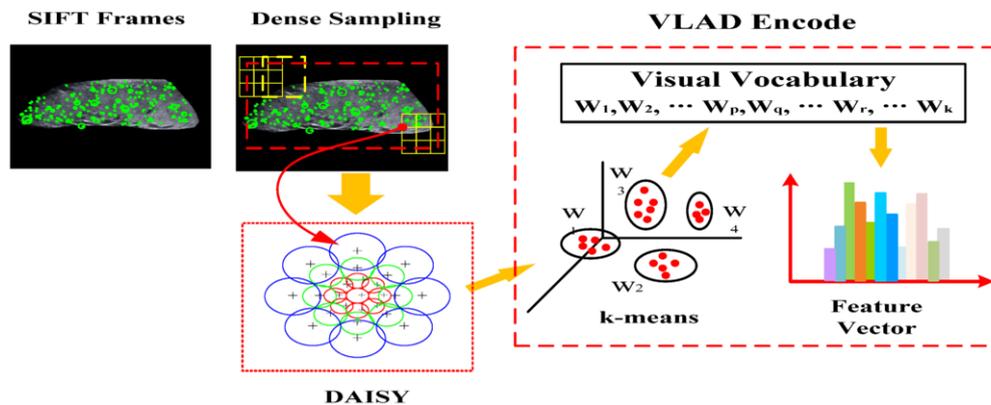

Fig. 1. System architecture.

Fig. 2. Construction of feature vector.

Besides, Gaussian mixture model (GMM) is conducted to produce k Gaussians using diagonal covariance matrix's assumptions. A single feature vector is formed based on the histogram representation. K-means is employed to cluster these DAISY features. Feature normalization is further utilized to boost performance. To avoid the problem of over-fitting, the staging decision rule is derived from the popularly applied support vector machine (SVM) classifier.

2.2. Dense sampling

For the interest point localization, traditional methods can only obtain the gradient point in the distributed space, and the fitting function is designed to locate the fixed interest point, location and scale of the gradient point. The fitting function can reduce the interest point with low contract and unstable one to enhance the stability. However, the detected interest point in the placental images by the sparse method remains too few, since the detected interest point cannot describe the point very well. Moreover, the local region is of different sizes. Since the placental images of different stages show small variations, it is very difficult to detect using the sparse descriptor. In general, the sparse descriptor has limited point and the detected point may have limited power to discriminate small differences. The accuracy will be decreased by the less discriminative features. By contrast, the dense descriptor extracts features from all the dense sampled points. In this paper, the dense sampling method in [10] is utilized. The neighboring and target region is sampled using a slide window to extract the features. Figure 3 illustrates the comparison of different descriptor and interest point detected by difference of Gaussian, Harris and Hessian and the dense sampling algorithms. It can be seen that the dense sampling method can attain better discriminability due to more samples captured instead of the unreliable captured interest point from the DoG, multiscale harris, and multiscale hessian methods.

2.3. DAISY

DAISY descriptor is first applied to wide base-line stereo matching. DAISY descriptor is motivated by SIFT and GLOH but shows better efficiency and performance. By graph-cuts, DAISY outperforms the window shape based competing SIFT local descriptor by dense map estimation algorithm. With Gaussian weighting and circularly symmetrical kernel, DAISY replaces SIFT weighted sums of gradient norms. It uses the convolution of gradients in a specific direction with several Gaussian filters to calculate the weighted sums at every pixel location. The histogram of per region is reused for all neighboring pixels, and the Gradient histogram of neighboring keypoints are also adopted to be scale and rotation invariant. The DAISY descriptor is circulated with the pixel point by three different circulars. Every circular with the same centering point computes the local gradient histogram; every pixel is sampled by 8 points by 45°. Every sampled point extracts the feature using the multiscale
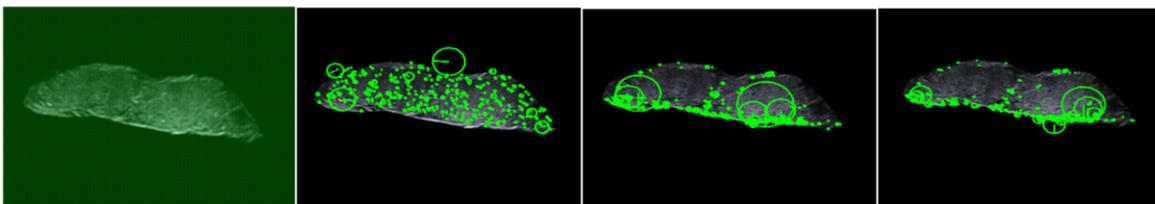

Fig. 3. Comparison of different descriptors and detectors based on the stage 3 placental image (from the left to right is the dense sampling method, DoG method, multiscale harris and multiscale hessian method).

technique to ensure its scale invariance. Different from the window shape in the traditional SIFT descriptor, the circular neighborhood configuration is adopted. Graph-cuts in DAISY descriptor use small image patches to guarantee spatial consistency by global optimization technique. Consequently, DAISY outperforms the traditional SIFT or GLOH region descriptors by reduced computational demand and simple correlation.

Let I be the input image, o is the orientation of the derivative, and $(\bullet)^+$ is the operator such that $(a)^+ = \max(a, 0)$, the orientation map $G_o = (\frac{\partial I}{\partial o})^+$, is first defined in DAISY descriptor. Every orientation map is convolved with Gaussian kernels of different Σ values several times to get the convolved orientation maps in different sized regions as $G_o^\Sigma = G_\Sigma * (\frac{\partial I}{\partial o})^+$ with G_Σ Gaussian kernel.

To extract features from DAISY descriptor, the descriptor is extracted in every pixel location, hence high feature dimensions are created in the DAISY descriptor. The high feature dimension makes application of placental maturity staging impractical. Since there is no keypoint detected in the DAISY descriptor, the dense sampling is capitalized on to have important point extraction.

For any point in (u, v), calculate the direction map in each direction as: $G_i, 1 \leq i \leq N$, the gradient histogram can be defined as:

$$\mathrm{h}_\Sigma(u, v) = \left[G_1^\Sigma(u, v), ..., G_H^\Sigma(u, v) \right]^T \tag{1}$$

and the histogram is normalized as: $\tilde{h}_\Sigma(u, v)$.

In DAISY descriptor, $D(u_0, v_0)$ in point (u_0, v_0) is represented by a series of the weighted vector $\tilde{h}_\Sigma(u, v)$, $\tilde{h}_{\Sigma_o}^T(1_j(u_0, v_0, \mathrm{R}_Q)$ is the corresponding histogram gradient for the Q circular loop in j sampling point. Every sampling point has the same Gaussian scale in every sampling point. In our experiment, the dimension of the DAISY descriptor is: $D = S \times H = 200$, where S is the number of the scales, and the H is the height of the features.

2.4. VLAD

The classification performance is improved first by feature encoding. The main idea of VLAD [8] is to calculate the difference in each cluster center. For a learned codebook by k-means: $\{\mu_k, k = 1, ..., K\}$, a set of local descriptors: $\{x_m, m = 1, ..., N\}$, the feature vector is extracted by assigning neighboring as below:

$$NN(x_m) = \underset{\mu_k}{\mathrm{argmin}} \|x_m - \mu_k\|. \tag{2}$$

v_k is calculated to form the feature vector by concatenation:

$$v_k = \sum_{x_m:NN(x_m)=\mu_k} x_m - \mu_k.$$ (3)

All feature vectors are normalized to form the final feature vectors for placental feature representation.

To better fit data by a GMM model, higher order statistics (i.e. derivative) are concatenated together. Gaussian means and variances of the first and second order derivatives [4] between features and GMM center are computed by:

$$\Phi_k^{(1)} = \frac{1}{N\sqrt{w_k}} \sum_{m=1}^{N} \gamma_m(k) \left(\frac{x_m - \mu_k}{\sigma_k} \right)$$ (4)

$$\Phi_k^{(2)} = \frac{1}{N\sqrt{2w_k}} \sum_{m=1}^{N} \gamma_m(k) \left(\frac{(x_m - \mu_k)^2}{\sigma_k} - 1 \right)$$ (5)

where $\{w_k, \mu_k, \sigma_k\}$ are the GMM mixture weights, means, and diagonal covariance. $\gamma_m(k)$ is the soft assignment weight of the m-th feature x_m of the k-th Gaussian. The derivatives of the log-likelihood of the GMM model are encoded by FV. The main purpose of the encoding is to discriminate the distribution difference between a specific test image and all fitted training image. FV ϕ is obtained by concatenating the difference vectors together:

$$\phi = [..., \Phi_1^{(1)}, \Phi_1^{(2)}, ..., \Phi_k^{(1)}, \Phi_k^{(2)}, ...]$$ (6)

PCA is performed to reduce the dimension of feature vector as well as processing time. Since the uncorrelated features and GMM covariance matrices of diagonal assumption are consistent, PCA whitening is also applied to ensure that diagonal covariance matrix assumption is satisfied.

3. Experimental results

In our experiment, there are a total of 311 placental images, 112 of which belongs to stage 0, 68 belongs to stage 1, 85 belongs to stage 2, and 36 belongs to stage 3. All images were acquired by a

Fig. 4 Image samples of 4 stages(from the left to right is stage 0 image, stage 1 image, stage 2 image, and stage 3 image).

commercial ultrasound scanner (Acuson Sequoia 512, Siemens Medical Solutions, USA) from Shenzhen Maternal and Child Health Hospital. Fetal gestational age ranges from 18 to 40 weeks. Conventional US sweep was performed to obtain images on pregnant women in the supine position by a radiologist with more than five years of experience in US obstetrics. The placental image samples of 4 classes are shown in Figure 4. Our system was implemented by the mixed programming technology of Matlab and C++. The interest detection and feature extraction time for an image (size: 800×700) is 6 seconds (32GBs RAM, double quad-core multithreaded server with a single CPU). The whole processing time for the testing step requires less than 1 second on a single CPU core.

   To evaluate the proposed method, quantitative measurements such as staging accuracy (the placental maturity is correctly staged but the actual number of samples are not in this stage), the mean average precision (mAP), sensitivity and specificity were utilized for performance evaluation. All experiments were repeated 10 times, and the average results are reported.

   The staging results of the ROC curves in terms of sensitivity and specificity are plotted in Figure 5. It can be seen that the highest staging performance is achieved by stage 0 since it is easy to separate it from other stages. Since stage 0 show almost perfect separation, only ROC curves of stage 1-3 are provided in this paper. It is known that stage 3 is often the most confusing one in placental maturity testing, and this is illustrated in Figure 6. Higher discriminative power is required for later stages since

(a) Stage1   (b) Stage 2   (c) Stage 3
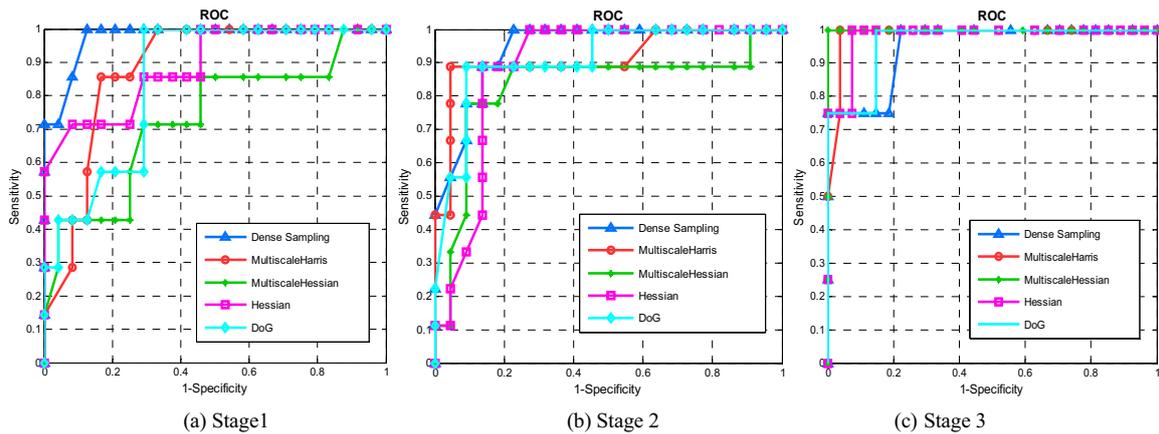
Fig. 5. ROC curves of the staging algorithm in different stages; (a) stage 1, (b) stage 2 and (c) stage 3.
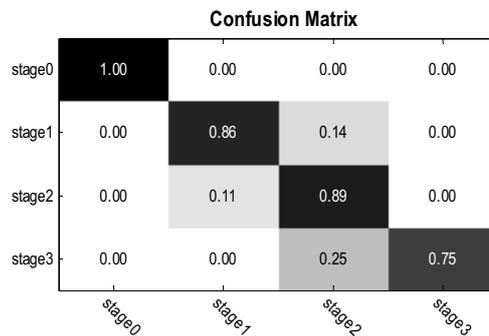

Fig. 6. Confusion matrix of placental maturity staging.

the training samples have higher variations. To obtain better performance, more discriminative learning by feature selection can be designed. It also can be observed that performance of the dense sampling outperforms other types of interest point detection method due to more samples generated for discrimination between different stages.

The confusion matrix of the placental staging is described in Figure 6. Rows in Figure 6 represent the actual staging level, whereas the columns are depicting the predicted staging level. The diagonal elements in the matrix represent the mean accuracy of each stage. As shown in the confusion matrix, the mean accuracy of 4 stages is 0.874, which is very high. In addition, the mean staging accuracy for each stage is also very high from the confusion matrix. Therefore, the promising performance achieved by the dense descriptor illustrates the effectiveness of the placental maturity staging.

Table 2 shows the classification accuracy, mAP, sensitivity and specificity results for the 4 different stages. It is observed that mAP is often higher than accuracy in the placental maturity staging. Compared with the traditional interest point methods such as multiscale hessian, hessian, DoG and mutiscale harris, dense sampling shows better performance than these traditional methods. The main reason is that the traditional method may lose the information during feature extraction of placental images. Though dense sampling may lose some information by the subsampling method, the extracted features are still effective for staging, hence contributing to the remarkable staging performance seen.

Table 3 shows staging results with different combination of feature descriptors and encoding methods. It is observed that DAISY feature descriptor outperforms the traditional feature descriptor such as SIFT. One explanation is that although DAISY feature have the same scale and translation invariant feature as SIFT, it can obtain more features from the sampled image. As a result, the performance is boosted by the dense sampled feature descriptor. Moreover, VLAD method is utilized to boost the performance. It is shown in Table 3 that VLAD outperforms the traditional BoVW method due to incorporation of higher statistics in the encoding method. Aggregating vectors contribute to better performance than the commonly used BoVW method. Introducing higher vectors can boost the performance not only in the context of placental image staging, but also in the other classification problems.

Table 2

Placental maturity grading results with different methods

Method	Accuracy	mAP	Sensitivity	Specificity
Multiscale Hessian	0.670	0.814	0.891	0.670
DoG	0.675	0.836	0.906	0.675
Hessian	0.669	0.863	0.907	0.669
Multiscale Harris	0.704	0.871	0.915	0.704
Dense Sampling	0.874	0.925	0.996	0.874

Table 3

Placental maturity grading results with feature descriptor and encoding method

Feature	Encode	Accuracy	mAP	Sensitivity	Specificity
SIFT	BoVW	0.776	0.893	0.943	0.776
DAISY	BoVW	0.767	0.864	0.934	0.767
SIFT	VLAD	0.818	0.942	0.946	0.818
DAISY	VLAD	0.874	0.925	0.996	0.874

4. Conclusion

In this paper, an automatic staging algorithm for placenta maturity is presented using the dense descriptor. Dense features have an advantage of extracting more discriminative features, where the aggregating vectors can boost the performance by introducing higher statistics method. Our experimental results showed promising performance in the context of placenta maturity staging. The staging algorithm is an effective way to provide assisted diagnosis for placental maturity staging. Future implementation of this method may reduce the physician's workload while reducing time-consuming visual surveillance in clinical practice.

Acknowledgement

References

[1]   R.T. Ribeiro, R.T. Marinho and J.M. Sanches, Classification and staging of chronic liver disease from multimodal data, Biomedical Engineering **60** (2013), 1336–1344.
[2]   Z. Liu, H. Zheng and S. Lin, Application of multi-classification support vector machine in the b-placenta image classification, Proc. of Int. Conf. on Computational Intelligence and Software Engineering, 2009, 1–4.
[3]   K. Mikolajczyk and C. Schmid, Scale & affine invariant interest point detectors, International Journal of Computer Vision **60** (2004), 63–86.
[4]   J. Sánchez, F. Perronnin, T. Mensink and J. Verbeek, Image classification with the fisher vector: Theory and practice, Int. J. of Comput. Vis. **105** (2013), 222–245.
[5]   Z. Wang, B. Fan and F. Wu, Local intensity order pattern for feature description, Proc. of Int. Conf. on Computer Vision, 2011, 603–610.
[6]   E. Tola, V. Lepetit and P. Fua, Daisy: An efficient dense descriptor applied to wide-baseline stereo, IEEE Trans. on Pattern Anal. and Mach. Intell. **32** (2010), 815–830.
[7]   S. Lazebnik, C. Schmid and J. Ponce, Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, Proc. of IEEE Conf. on Comput. Vis. and Pattern Recognit. **2** (2006), 2169–2178.
[8]   H. Jégou, F. Perronnin, M. Douze and C. Schmid, Aggregating local image descriptors into compact codes, IEEE Trans. on Pattern Anal. and Mach. Intell. **34** (2012), 1704–1716.
[9]   P.A. Linares, P.J. McCullagh, N.D. Black and J. Dornan, Feature selection for the characterization of ultrasonic images of the placenta using texture classification, Proc. of Int. Symp. on Biomed. Imag. **2** (2004), 1147–1150.
[10]   A. Vedaldi and B. Fulkerson, VLFeat: An open and portable library of computer vision algorithms, Proc. of Int. Conf. on Multimedia, 2010, 1469–1472.