Automated Macular Pathology Diagnosis in Retinal OCT Images Using Multi-Scale Spatial Pyramid with Local Binary Patterns

Yu-Ying Liu[1], Mei Chen[2], Hiroshi Ishikawa[3,4], Gadi Wollstein[3], Joel S. Schuman[3,4], and James M. Rehg[1]

[1] School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA
[2] Intel Labs Pittsburgh, Pittsburgh, PA
[3] UPMC Eye Center, University of Pittsburgh Medical Center, Pittsburgh, PA
[4] Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA

Abstract. We address a novel problem domain in the analysis of optical coherence tomography (OCT) images: the diagnosis of multiple macular pathologies in retinal OCT images. The goal is to identify the presence of normal macula and each of three types of macular pathologies, namely, macular hole, macular edema, and age-related macular degeneration, in the OCT slice centered at the fovea. We use a machine learning approach based on global image descriptors formed from a multi-scale spatial pyramid. Our local descriptors are dimension-reduced Local Binary Pattern histograms, which are capable of encoding texture information from OCT images of the retina. Our representation operates at multiple spatial scales and granularities, leading to robust performance. We use 2-class Support Vector Machine classifiers to identify the presence of normal macula and each of the three pathologies. We conducted extensive experiments on a large dataset consisting of 326 OCT scans from 136 patients. The results show that the proposed method is very effective.

1 Introduction

Optical Coherence Tomography (OCT) is a non-contact, non-invasive 3-D imaging technique which performs optical sectioning at micron resolution. It has been widely adopted as the standard of care in ophthalmology for identifying the presence of disease and its progression [1]. This technology measures the optical back scattering of the tissues, making it possible to visualize intraocular structures, such as the macula and the optic nerve. An example 3D ocular OCT scan is given in Fig. 1a. The ability to visualize the internal structures of the retina (the z-axis direction in Fig. 1a) makes it possible to diagnose diseases, such as glaucoma and macular hole, objectively and quantitatively.

Although OCT imaging technology continues to evolve, technology for automated OCT image analysis and interpretation has not kept pace. With OCT data being generated at increasingly higher sampling rates, there is a strong need for automated analysis to support disease screening and diagnosis. This need is amplified by the fact that an ophthalmologist making a diagnosis under
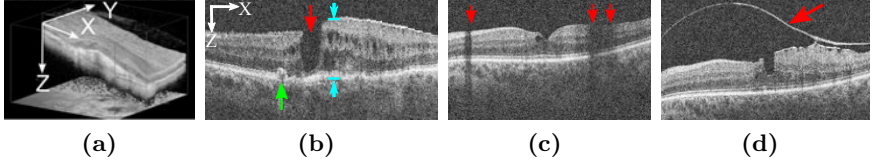
Fig. 1. (a) Example 3D macular OCT scan. (b)(c)(d) Example x-z slice with (b) MH (red), ME (blue), and AMD (green), (c) shadowing effects, (d) a detached tissue.
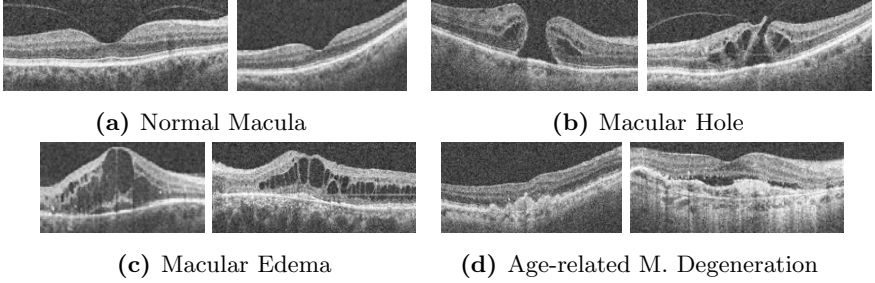

(a) Normal Macula   **(b)** Macular Hole

(c) Macular Edema   **(d)** Age-related M. Degeneration

Fig. 2. Characteristics of (a) normal macula: a smooth depression shows at the center (fovea), (b) macular hole (MH): a partial or full hole formation at the fovea, (c) macular edema (ME) : retinal thickening and liquid accumulation appears as black blobs around the fovea, (d) age-related macular degeneration (AMD): irregular contours usually extruded in dome shapes appear at the bottom layer of the retina.

standard clinical conditions does not have the assistance of a radiologist in interpreting OCT data. There have been several prior works addressing topics in ocular OCT image processing, such as intra-retinal layer segmentation [2] and local quality assessment [3]. However, to our knowledge, *there has been no prior work on automated macular pathology identification in OCT images.*

The macula is located at the center of the retina and is responsible for highly-sensitive, accurate vision. Acute maculopathy can lead to the loss of central vision and even blindness. For example, a particular type of maculopathy, called age-related macular degeneration (AMD), is the leading cause of visual loss among elderly persons. The Beaver Dam Eye Study reported that 30% of individuals aged 75 and older have some form of AMD, which has increasingly important social and economic impact in the United States [4]. The screening and diagnosis of maculopathy is a substantial public health problem.

We present a method for automatically identifying macular pathologies given a manually-selected x-z 2D slice centered at the fovea in a retinal OCT scan. Specifically, we identify the presence of the normal macula and each of the following macular pathologies, macular hole (MH)[1], macular edema (ME), and AMD.

---

[1] For simplicity, we include both macular hole and macular pseudohole in MH category.

Fig. 2 gives example images for each pathology. Note that multiple pathologies can coexist in one eye, as depicted in Fig. 1b. In this case, the automated identification tool should report the existence of each pathology.

2 Approach

Automated pathology identification in ocular OCT images is complicated by four factors. First, the co-existence of pathologies can complicate the overall appearance (see Fig.1b). Second, the measurement of reflectivity of the tissue is affected by the optical properties of the overlaying tissues [1], e.g, opaque media or the high rate of light absorption by blood vessels will produce shadowing effects (see Fig.1c). Third, a variety of proliferated or detached tissues can appear, e.g. membrane or floaters (see Fig.1d). Fourth, a portion of the image may have low quality due to imperfect imaging [3]. As a result of this high variability, attempts to hand-craft a set of features or rules to identify each pathology are unlikely to succeed. Instead, we propose to use a pattern-based global image representation combined with machine learning techniques to automatically discover the discriminative patterns for each pathology from training examples.

Our method consists of the following three steps, which are illustrated in Fig. 3. First, image alignment is performed to reduce the appearance variation across scans. Second, we construct a global descriptor for the aligned image by using multi-scale spatial pyramid (MSSP) and the dimension-reduced Local Binary Pattern (LBP) histogram [5] based on Principle Component Analysis (PCA), in order to represent the retina at different spatial scales and granularities. Finally, for each pathology, we train a 2-class non-linear Support Vector Machine (SVM) [6] with radial basis function (RBF) kernel using a labeled training set. We now describe each step in detail.

2.1 Retina Alignment

Since the imaged retinas have large variations in their inclination angles, positions, and natural curvatures across scans, as shown in Fig.2, it is desirable
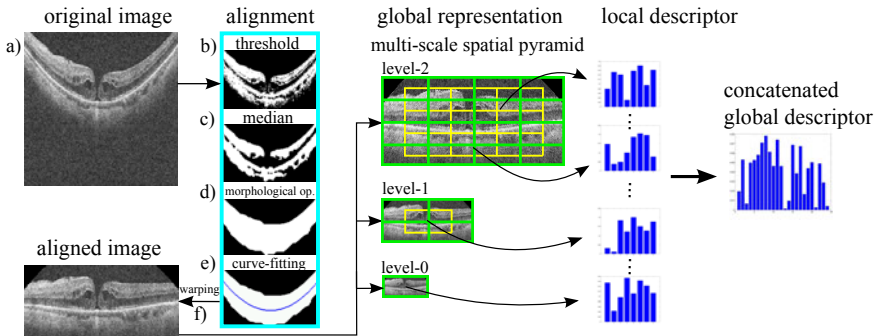

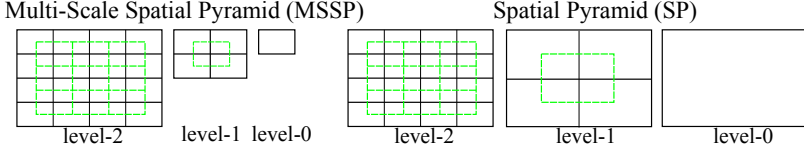**Fig. 3.** Stages in the constructing of the global image descriptor

Fig. 4. Structures of a 3-level MSSP and SP. The green lines indicate the overlapped blocks which can also be added in the global representation.

to roughly align the retinas to reduce these variations before constructing our feature representation. To this end, we use a simple procedure as follows: (1) threshold the original image (Fig 3, a) to detect most of the retina structures (Fig 3, b); then, apply a median filter to remove noise and thin detached tissues (Fig 3, c); (2) find the entire retina by using morphological closing and then opening; by closing, we fill-up black blobs inside the retina, and by opening, to remove thin or small objects outside the retina (Fig 3, d); (3) fit the found retina with a quadratic curve by using least-square curve fitting (Fig 3, e); (4) warp the retina to be approximately horizontal by translating each column according to the fitted curve, and crop it in the z-direction (Fig 3, f).

2.2   Multi-Scale Spatial Pyramid (MSSP)

There are three motivations for our choice of a global spatially-distributed feature representation for OCT imagery based on MSSP. First, pathologies are often localized to specific retinal areas, making it important to encode spatial location. Second, the context provided by the overall appearance of the retina is important for correct interpretation; e.g., in Fig. 1c, we can distinguish between a shadow and macular hole only in the context of the entire slice. Third, pathologies can exhibit discriminating characteristics at both small and large scales. Thus, both micro-patterns and macro-patterns should be represented. For these reasons, we use a global image representation which preserves spatial organization in conjunction with multi-scale modeling.

We propose to use the multi-scale version of spatial pyramid (SP) [7], denoted as MSSP, to capture the geometry of the aligned retina at multiple scales and spatial resolutions. This global framework, MSSP, was recently proposed in [8], where it was successfully applied to challenging scene classification tasks. For clarity, in Fig.4, we illustrate the differences between a 3-level MSSP[2] and SP frameworks. The local features computed from all spatial blocks are concatenated in a predefined order to form an overall global descriptor, as illustrated in Fig. 3. Note that we also add the features from the overlapped blocks (the green blocks in Fig. 4) to reduce boundary effects.

---

[2] To form a k-level MSSP, for each level l ($0 \leq l \leq (k-1)$), we rescale the original image by 2^{l-k+1} using bilinear interpolation, and divide the rescaled image into 2^l blocks in each dimension.

2.3   Histogram of LBP and Dimensionality Reduction

Local binary pattern (LBP) [5] is a non-parametric kernel which summarizes the local structure around a pixel. LBP is known to be a highly discriminative operator and has been successfully applied to several tasks in computer vision [9,8] and medical imaging [10,11].

While there are several types of LBP, we follow [9,10,11] in adopting $LBP_{8,1}$ to capture the micro-patterns that reside in each local block. The $LBP_{8,1}$ operator derives an 8 bit binary code by comparing the center pixel to each of its 8 nearest neighbors in a 3×3 neighborhood (see Fig. 5a). The resulting 8 bits are concatenated circularly to form an LBP code in the range [0 255]. For each block of pixels in the MSSP, we compute the histogram of LBP codes to encode the statistical distribution of different micro-patterns, such as spots, edges, corners, and flat areas [5]. Histogram descriptors have proven to be an effective means to aggregate local intensity patterns into global discriminative features. In particular, they avoid the need to precisely localize discriminating image structures, which can be difficult to do in complex and highly variable OCT images. Since we compute LBP histogram in both original and rescaled image blocks, the distribution of both micro-patterns and macro-patterns can be encoded. Note that many previous works applied LBP only in the original image [9,10,11], which may not capture the large-scale patterns effectively.

Although a single $LBP_{8,1}$ histogram has only 256 bins, the concatenation of histograms from each block to form the global feature vector results in an impractically high dimension. We adopt PCA to reduce the dimension of LBP histogram, as proposed in [8]. We denote it as $LBP_{8,1}^{pca}$.

It is important to note that previous works [9,10] employing LBP histogram features have adopted an alternative approach to dimensionality reduction, called uniform LBP [5], which we have found to be less effective than $LBP_{8,1}^{pca}$ for AMD category. An LBP pattern is called *uniform* if it contains at most two bitwise transitions, as demonstrated in Fig. 5b. A histogram of $LBP_{8,1}^{u2}$ is formed by retaining occurrences of each of 58 uniform patterns and putting all occurrences of 198 non-uniform patterns, denoted as $LBP_{8,1}^{\neg u2}$, to a single bin, resulting in 59 bins in total. It was observed in [5] that $LBP_{8,1}^{u2}$ occupied 90% of all $LBP_{8,1}$ patterns in pixel count, when computed from image textures; however, as recently noted in [12], when LBP codes are computed in the rescaled images, $LBP_{8,1}^{u2}$ may no longer be in the majority. More importantly, the distribution of individual $LBP_{8,1}^{\neg u2}$ patterns can contain important distinctive information for category discrimination, despite of its low counts (see Fig. 6). To advocate the use of PCA for LBP dimension reduction, we will compare the performance of using $LBP_{8,1}^{pca}$, $LBP_{8,1}^{u2}$, and $LBP_{8,1}$ in Section 3.

2.4   Classification Using Support Vector Machine (SVM)

After computing global features, we train a 2-class non-linear SVM with RBF kernel and probabilistic output [6] for each pathology using a 1 vs. the rest approach. The probability scores for each classifier are compared to a set of decision thresholds to determine the corresponding sensitivities and specificities.
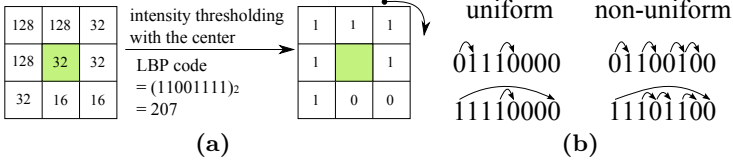
Fig. 5. Examples of (a) $LBP_{8,1}$ code computation, (b) uniform and non-uniform LBP
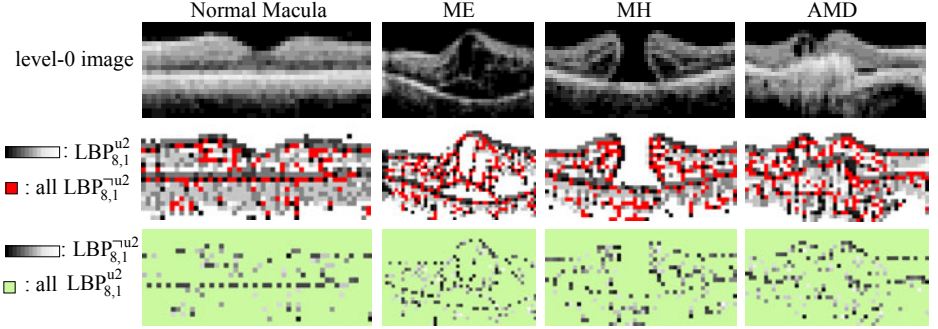

Fig. 6. Visualization of $LBP_{8,1}^{u2}$ and $LBP_{8,1}^{\neg u2}$ on example level-0 images. All $LBP_{8,1}^{\neg u2}$ are shown in red in the 2nd row, and the individual $LBP_{8,1}^{\neg u2}$ codes are shown in gray-level in the 3rd row. We find that although $LBP_{8,1}^{\neg u2}$ patterns have low counts, most of them reside in the important contours and can therefore be useful for discrimination.

3   Experimental Results

We collected a dataset consisting of 326 macular spectral-domain OCT scans from 136 patients. The original resolution of the scans was either 200x200x1024 or 512x128x1024 scanned in 6x6x2 mm volume in width(x), height(y) and depth(z). We rescaled all x-z slices to 200x200. For each scan, the x-z slice centered at the fovea and the pathologies it contained was identified independently by two expert ophthalmologists. Among all 326 scans, 96.2% of the selected fovea-centered slices from the two experts were within 4 slice distance; for these scans, the labeling agreement for normal macula, ME, MH, AMD is 98.6%, 91.7%, 96.5%, 92.0%, at the patient-level respectively. This imperfect agreement provides further evidence of the need for an objective computer-aided diagnosis tool. We used the labels from one ophthalmologist as the ground truth.

The statistics of the labeled dataset are as follows. We have 67 positive and 259 negative scans for normal macula, 205 and 121 for ME, 81 and 245 for MH, 103 and 223 for AMD. From the standpoint of patient statistics, we have 57 positive and 79 negative patients for normal macula, 87 and 49 for ME, 34 and 102 for MH, 36 and 100 for AMD. We used 10-fold cross validation at the patient level; i.e., 10% of positive and negative patients were put in the testing

set for each run. Here, for a specific pathology, the patients that have at least one positive scan were counted as positive patients; also, all images from the same patient were put together in either the training or testing set. We used one fold of training data to do SVM model selection based on the area under the receiver operator characteristic curve (AUC) performance, and fixed the found best parameters throughout all the experiments. We normalized the sensitivity and specificity values across patients and used AUC as the comparison metric.

To validate the performance of LBP histograms as features, we compared them to several other popular descriptors, including mean and standard deviation of intensity (M,S), intensity histogram (I), and orientation histogram (O). Each feature is employed with a 3-level MSSP with overlapped blocks[3]. The orientation histogram [13] is formed from the gradient direction and magnitude computed from 2×2 neighborhoods. For I and O features, we used the quantization of 32 bins in intensity and angle, respectively, since this produced the best results. For LBP histogram, we quantize the intensity image to 32 levels before LBP encoding in order to suppress pixel noise. This quantization improves the accuracy of LBP by about 0.8% on average. For $LBP_{8,1}^{pca}$ computation, the principle axes were derived from the training images of each fold separately.

Overall, from Table 1, $LBP_{8,1}^{pca}(32)$ achieved the best average performance. In details, we find for normal macula and ME, the AUC are above 0.95 for almost all descriptors; this can be attributed to the reduced variation in appearance for the normal category and to the large amount of positive training data for ME. The AUC results of MH and AMD are lower, which can be due to their greater appearance variation and lack of data. In these two categories, LBP with dimension reduction ($LBP_{8,1}^{u2}$ and $LBP_{8,1}^{pca}$) outperforms the other descriptors by a clear margin. Note that the use of all 256 bins of LBP histogram gives the worst results, presumably due to overfitting in the high dimensional feature space.

We then compare the results of $LBP_{8,1}^{u2}$ and $LBP_{8,1}^{pca}$. From Table 1, $LBP_{8,1}^{u2}$ and $LBP_{8,1}^{pca}$ have similar performance for the first three categories, where the pathologies are described mostly by smooth or global shape changes. In AMD category, however, $LBP_{8,1}^{pca}(32)$ is clearly superior to $LBP_{8,1}^{u2}$. Our conjecture is that since AMD is characterized by the irregular contours of the bottom retinal layer, the removal of individual non-uniform patterns, as used in $LBP_{8,1}^{u2}$, can result in the loss of important discriminative information. Finally, the use of first 32 principal components seems sufficient for category discrimination.

In Table 2, we compare the performance of 3-level multi-scale spatial pyramid (MSSP) with a 3-level spatial pyramid (SP) and single level-2 spatial division (SL)[4]. MSSP achieves the best overall performance; in AMD category, MSSP outperforms the other frameworks with a large margin (0.888 vs. 0.826). When features from the overlapped blocks are also used (denoted as "+O"), the performance of all frameworks is improved.

---

[3] If the feature dimension of the block descriptor is d, a 3-level MSSP with the overlapped blocks will result in a global descriptor of length $d \times 31$, where 31 is the total number of blocks.

[4] Only $4 \times 4 = 16$ spatial blocks derived from the original image were used.

Table 1. The AUC results obtained from using different local block descriptors. Number in parenthesis represents the feature dimension of each local descriptor.

Pa./Feature	M,S(2)	I(32)	O(32)	$LBP_{8,1}^{u2}(59)$	$LBP_{8,1}^{pca}(32)$	$LBP_{8,1}^{pca}(59)$	$LBP_{8,1}(256)$
Normal M.	0.965	0.970	0.983	**0.991**	0.987	**0.991**	0.931
ME	0.951	0.963	0.958	**0.965**	0.962	0.962	0.845
MH	0.714	0.826	0.845	**0.901**	0.894	0.879	0.774
AMD	0.784	0.824	0.857	0.867	**0.888**	0.885	0.693
Ave. of all	0.854	0.895	0.911	0.931	**0.933**	0.929	0.811
Ave. of MH,AMD	0.749	0.825	0.851	0.884	**0.891**	0.882	0.734

Table 2. The AUC results obtained from employing different global frameworks with $LBP_{8,1}^{pca}(32)$ local descriptors. "+O": add the overlapped blocks.

Pa./Framework	MSSP	SP	SL	MSSP+O	SP+O	SL+O
Normal M.	0.985	0.979	0.984	**0.987**	0.984	0.987
ME	0.957	0.949	0.951	**0.962**	0.960	0.961
MH	0.891	0.880	0.872	0.894	**0.895**	0.893
AMD	0.888	0.826	0.826	**0.888**	0.849	0.843
Ave. of all	0.930	0.909	0.908	**0.933**	0.922	0.921

4   Conclusion and Future work

In this paper, we propose an effective approach to diagnose multiple macular pathologies in retinal OCT images. First, we align the slices to reduce their appearance variations. Then, we construct our global image descriptor by using multi-scale spatial pyramid (MSSP), combined with dimension-reduced LBP histogram based on PCA as the local descriptors. This approach encodes both the geometry and textures of the retina in a principled way. A binary non-linear SVM classifier is trained for each pathology to identify its presence. We evaluate our approach by comparing its performance with that of other popular global representation and local descriptors. Our results demonstrate the effectiveness and validity of the proposed approach.

There are several future directions we plan to explore. First, we will investigate the effects of including other complimentary features, e.g., encoding shape features from the edge image, so that more discriminative representation can be constructed. Second, we would like to apply a similar method to automatically localize the slice of the anatomic landmark (e.g. fovea) in the 3D scan. Third, we plan to extend our approach to additional retinal pathologies.

References

1. Schuman, J.S., Puliafito, C.A., Fujimoto, J.G.: Optical Coherence Tomography of Ocular Diseases, 2nd edn. (2004)
2. Garvin, M.K., Abràmoff, M.D., Kardon, R., Russell, S.R., Wu, X., Sonka, M.: Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search. TMI 27(10), 1495 (2008)
3. Barnum, P., Chen, M., Ishikawa, H., Wollstein, G., Schuman, J.: Local quality assessment for optical coherence tomography. In: ISBI (2008)
4. Klein, R., Klein, B.E.K., Linton, K.L.P., DeMets, D.L.: The Beaver Dam Eye Study: The relation of age-related maculopathy to smoking. American Journal of Epidemiology 137(2), 190–200 (1993)
5. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. on Pattern Analysis and Machine Intelligence(TPAMI) 24(7), 971 (2002)
6. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines (2001)
7. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: CVPR (2006)
8. Wu, J., Rehg, J.M.: Where am I: Place instance and category recognition using spatial PACT. In: CVPR (2008)
9. Ahonen, T., Hadid, A., Pietikäinen, M.: Face description with local binary patterns: Application to face recognition. TPAMI 28(12), 2037 (2006)
10. Oliver, A., Lladó, X., Freixenet, J., Martí, J.: False positive reduction in mammographic mass detection using local binary patterns. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 286–293. Springer, Heidelberg (2007)
11. Sørensen, L., Shaker, S.B., de Bruijne, M.: newblock Texture classification in lung CT using local binary patterns. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 934–941. Springer, Heidelberg (2008)
12. Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S.Z.: Learning multi-scale block local binary patterns for face recognition. In: Int'l. Conf. on Biometrics (2008)
13. Freeman, W.T., Roth, M.: Orientation histogram for hand gesture recognition. In: Intl. Workshop on Automatic Face and Gesture Recognition (1995)