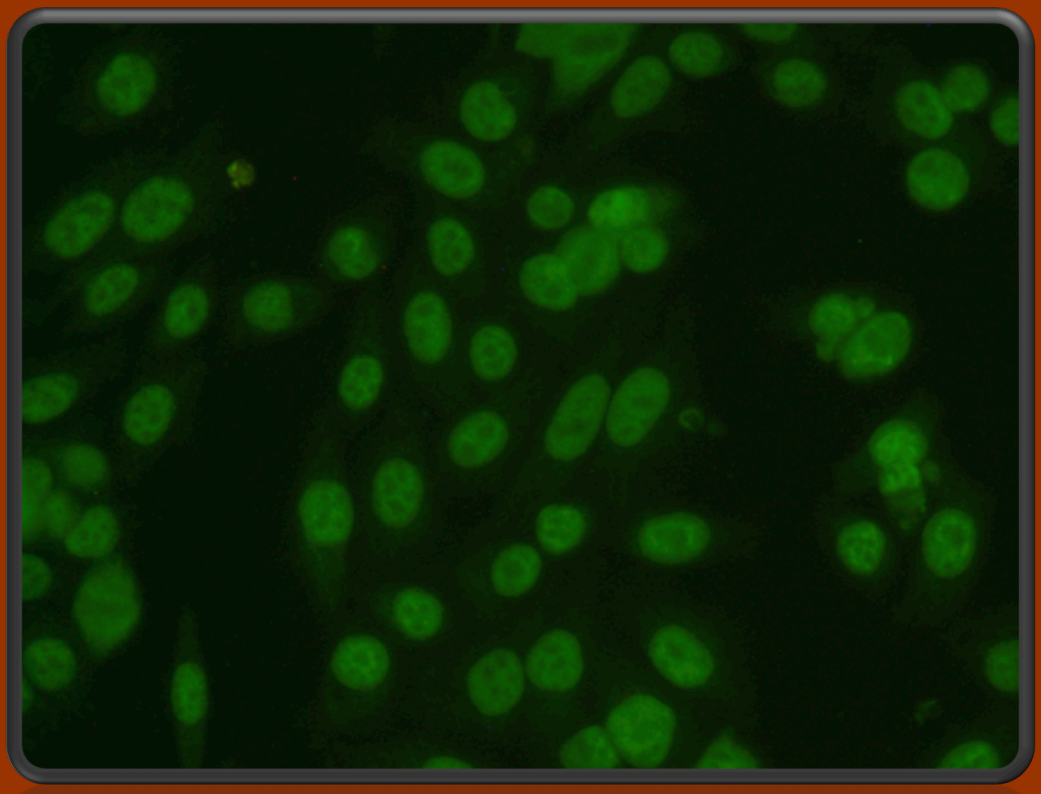
Performance Evaluation of Indirect Immunofluorescence Image Analysis Systems

hosted by the
22nd International Conference on Pattern Recognition (ICPR 2014)



Organizers

Brian C. Lovell - *(lovell@itee.uq.edu.au, University of Queensland, Australia)*
Gennaro Percannella - *(pergen@unisa.it, University of Salerno, Italy)*
Mario Vento - *(mvento@unisa.it, University of Salerno, Italy)*
Arnold Wiliem - *(a.wiliem@uq.edu.au, University of Queensland, Australia)*

http://i3a2014.unisa.it/

Performance Evaluation of Indirect Immunofluorescence Image Analysis Systems

Brian C. Lovell*, Gennaro Percannella†, Mario Vento† and Arnold Wiliem*
*University of Queensland, Australia
†University of Salerno, Italy

I. Description of the initiative

In the recent years we have assisted to a progressively growing number of applicative areas of Pattern Recognition, mainly devoted to the exploitation of cutting edge scientific methodologies for the solution of problems of relevant interest to civil society. This trend is generating new communities, as aggregation of scientists having as common aim the development of systems more or less prompt to be transformed into real working prototypes.

In the field of medical image analysis this trend has been even more evident than in others, as the availability of assisted diagnosis tools would allow the medical community to increase their productivity jointly with an improvement of the quality and precision of the diagnostic act.

Among all, rather novel interests are concentrating on the indirect immunofluorescence images (IIF), i.e. images obtained by making biological tissue interacting with special sources of light, so as to generate fluorescent image responses; these revealed especially suited for searching antibodies in the patient serum, allowing to detect the presence of autoimmune diseases. Due to its effectiveness, diagnostic tests for systemic autoimmune diseases are now becoming more and more interesting to both academic and industrial communities.

This research topic is achieving new enthusiasm and interest among scientists and the size of the community may now be considered significant. An international contest, organized by researchers above at the 21st International Conference on Pattern Recognition (ICPR2012) and entitled HEp-2 cell classification (http://mivia.unisa.it/hep2contest), is a concrete and recent demonstration of the presence of a very active scientific community: over 100 registrations and 28 final submissions of executable code were registered, making it the most popular contest at ICPR 2012 [1]. This interest has been further confirmed by the significant number of registrations and submissions (more than 100 registrations with 14 executable submissions) to the Competition on Cells Classification by Fluorescent Image Analysis (http://nerone.diiie.unisa.it/contest-icip-2013) hosted by the 20th IEEE International Conference on Image Processing (ICIP 2013).

The Performance Evaluation of Indirect Immunofluorescence Image Analysis Systems (http://i3a2014.unisa.it/) held at the 22nd International Conference on Pattern Recognition (ICPR 2014) is a benchmarking activity aimed at guaranteeing the continuity of this series of events and to allow to enlarge the community of researchers in this field.

The Performance Evaluation of Indirect Immunofluorescence Image Analysis Systems is an initiative jointly organized by the University of Salerno (Italy) and the University of Queensland (Australia) with the support of the Sullivan Nicolaides Pathology (SNP) - Queensland Medical Testing Laboratory, (Australia). The involvement of the SNP, a relevant laboratory that offers high qualified pathology services for doctors, private hospitals and nursing homes in Queensland northern New South Wales and Darwin, allowed to obtain significant and valuable datasets used for the contest.

In this edition of the competition, two different recognition tasks were proposed to the participants. In the remainder of this document we will refer to tham as Task 1 and Task 2. Each research team was allowed to apply to both tasks or just one. A detailed description of each task is proposed in the following sections.

II. Task 1: cell level classification

This classification task is a reproposition of the competition on Cells Classification by Fluorescent Image Analysis hosted on last September 2013 by the 20th IEEE International Conference on Image Processing (ICIP 2013). The main aim was to encourage the participation on one hand of new research teams that intended to start with a task that is simpler than the Task 2 described below and on the other hand of experienced teams that already competed in 2013 and intended to improve their algorithms. More in general, the idea is to maintain this task as a permanent competition with the aim of constantly monitoring the advances of the research in this field.

A. Problem formulation

Each participant is required to develop a classifier φ which classifies a set of HEp-2 cell images. Each image is represented by three-tuple $(\boldsymbol{I}, \boldsymbol{M}, \delta)$ [3], where:

- $\boldsymbol{I}$ represents the cell fluorescence image;
- $\boldsymbol{M}$ is the cell mask which is automatically extracted;
- δ represents the cell positivity strength which has two values weak/borderline or strong.

Let $\boldsymbol{Y}$ be a probe image, ℓ be its class label and $\mathcal{G} = \{(\boldsymbol{I}, \boldsymbol{M}, \delta)_1, ...(\boldsymbol{I}, \boldsymbol{M}, \delta)_n\}$ be a given gallery set. The classifier's task is to predict the probe label, $\hat{\ell}$. In other words, $\varphi : \boldsymbol{Y} \times \mathcal{G} \mapsto \hat{\ell}$, where ideally $\hat{\ell} = \ell$.

B. Dataset

The dataset has been collected between 2011 and 2013 at Sullivan Nicolaides Pathology laboratory, Australia. It utilizes

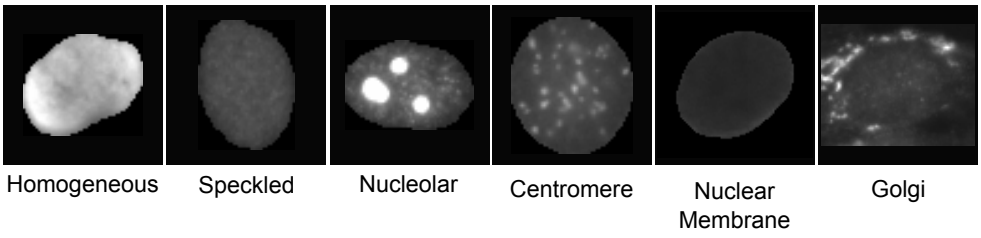


Fig. 1. Sample images from ICIP2013 dataset used for the competition of the Task 1.

419 patient positive sera, which were prepared on the 18-well slide of HEP-2000 IIF assay from Immuno Concepts N.A. Ltd. with screening dilution 1:80. The specimens were then automatically photographed using a monochrome high dynamic range cooled microscopy camera which was fitted on a microscope with a plan-Apochromat 20x/0.8 objective lens and an LED illumination source. Approximately 100-200 cell images were extracted from each patient serum. In total there were 68,429 cell images extracted: 13,596 images used for training, made available to the teams, and 54,833 for testing, privately mantained by the organizers.

The dataset contains the following six classes (see Fig. 1 for some examples) [2]:

- *homogeneous*: a uniform diffuse fluorescence covering the entire nucleoplasm sometimes accentuated in the nuclear periphery;
- *speckled*: these patterns have two sub-categories [1]:
  - *coarse speckled*: densely distributed, variously sized speckles, generally associated with larger speckles, throughout nucleoplasm of interphase cells; nucleoli are negative;
  - *fine speckled*: fine speckled staining in a uniform distribution, sometimes very dense so that an almost homogeneous pattern is attained; nucleoli may be positive or negative;
- *nucleolar*: brightly clustered large granules corresponding to decoration of the fibrillar centers of the nucleoli as well as the coiled bodies;
- *centromere*: rather uniform discrete speckles located throughout the entire nucleus;
- *golgi*: staining of a polar organelle adjacent to and paratly surrounding the nucleus, composed of irregular large granules. Nuclei and nucleoli are negative. Diffuse staining of the cytoplasm of dividing cells sometimes with accentuation around chromosomal material;
- *nuclear membrane*: a smooth homogeneous ring-like fluorescence of the nuclear membrane in interphase cells.

The labeling process involved at least two scientists who read each patient specimen under a microscope. A third expert's opinion was sought to adjudicate any discrepancy between the two opinions. We used each specimen label for the ground-truth of cells extracted from it. Furthermore, all the labels were validated by using secondary tests such as ENA and anti-ds-DNA in order to confirm the presence and/absence of specific patterns. Each cell image contained in the database is annotated with the following information:

- Cell pattern (one of the patterns defined above)
- Cell intensity
- Cell mask
- ID of the image which the cell belongs to

[1]In this dataset, we consider these two sub-categories as one category.

C. Evaluation

As there is distribution imbalance between the common and lesser common patterns, it is important to define an evaluation protocol which is not biased towards the common patterns. For this reason in this edition of the competition the mean class accuracy (MCA) is adopted as the measure of the methods' performance instead of the overall accuracy used at ICPR 2012 and ICIP 2013 contests.

Technically, we compute the accuracy for each pattern class and then we take the average of the per-class accuracies. In particular, let $\mathrm{CCR_k}$ be the correct classification rate for class k determined as follows.

$$\mathrm{CCR}_k = \frac{1}{N_k}\left(\mathrm{TP_k} + \mathrm{TN_k}\right) \tag{1}$$

where $\mathrm{TP_k}$ and $\mathrm{TN_k}$ are the number of true positive and true negative results on class k, respectively; N_k is the number of images that belong to class k. The mean class accuracy MCA is determined by:

$$\mathrm{MCA} = \frac{1}{K}\sum_{k=1}^{K}\mathrm{CCR}_k \tag{2}$$

where K is the number of pattern classes (*i.e.* K equals seven in this instance).

III. Task 2: specimen level classification

With this task, the scientific community had to face with a new and more challenging pattern recognition problem, i.e. the classification at specimen level. To this aim it has been used a brand new very large dataset of HEp-2 images collected with the support of the Sullivan Nicolaides Pathology testing laboratory. The dataset is composed by specimen images, each one containing a variable number of HEp-2 cells.

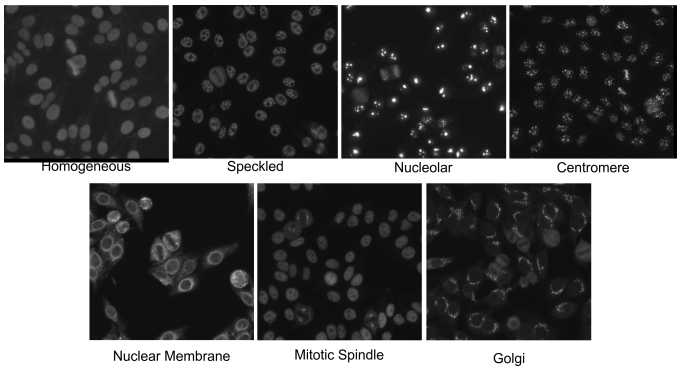


Fig. 2. Sample specimen images used for the Task 2.

A. Problem formulation

We use the style of [3] to describe the ANA IIF HEp-2 image classification problem, or ANA image classification problem in short. An ANA image $\boldsymbol{I}$ is represented by the three-tuple $\{\boldsymbol{I}, \boldsymbol{M}, \delta\}$ which consists of: **(i)** the Fluorescein Isothiocyanate (FITC) image channel which carries pattern information $\boldsymbol{I}$; **(ii)** a binary cell mask image $\boldsymbol{M}$ which are extracted from the *4',6-diamidino-2-phenylindole* (DAPI) image channel [2]; **(iii)** the fluorescence intensity $\delta = \{\text{weak}, \text{strong}\}$. The goal is to construct a classifier which classifies an image into one of the known ANA classes.

B. Dataset

The dataset was acquired in 2013 at Sullivan Nicolaides Pathology Laboratory, Australia. It was collected from 1001 patient sera with positive ANA test. Each patient sera were diluted to 1:80 and the specimen was photographed using a monochrome camera fitted on a microscope. Each specimen was photographed at four different locations, so that each specimen has four images. There are seven patterns considered: homogeneous, speckled, nucleolar, centromere, golgi aparatus (Golgi), nuclear membrane (NuMem) and mitotic spindle. As depicted in Fig. 2, the first four patterns are common ANA patterns and the last three are much lesser common or rare. In particular, with respect to the Task 1, here we considered also the mitotic spindle class for which we use the standard textual description proposed in [2]:

- *mitotic spindle* staining only in the triangular or "banana-shaped" pole area of the mitotic spindle in the metaphase cells

We deliberately combined both common and uncommon patterns to simulate the real scenario where pathology laboratories need to deal with both patterns. We divided the images into two sets: 252 images for training and 749 for testing.

The labelling process has involved at least two scientists who read each patient specimen under a microscope. A third expert's opinion has been sought to adjudicate any discrepancy between the two opinions. Each slide image is also provided with the corresponding class (one among the patterns defined above). Furthermore, all the labels were validated by using secondary tests such as ENA and anti-ds-DNA in order to confirm the presence of specific patterns.

Each specimen image contained in the database is annotated with the following information:

- Staining pattern of the specimen
- Intensity of the specimen
- Mask of the specimen

C. Evaluation

Submissions are evaluated using the same performance index adopted for Task 1 (Mean Class Accuracy).

IV. Participation

The competition received more than 100 registrations from around the world with 11 submissions for Task 1 and 7 submissions for Task 2. In the following we report the members and affiliations of all the teams that participated to the Competition on Cells Classification by Fluorescent Image Analysis. We refer to each method of each team using the name of the corresponding author of the submission. Furthermore, for each group the participation to the task 1 and/or 2 is indicated using **T1** and/or **T2**.

- **CODRESCU (T1)**: C. Codrescu, *University of Salzburg, Austria*.
- **ENSAFI (T1 and T2)**: S. Ensafi[1], S. Lu[2], A. Kassim[1], C.L. Tan[1], [1]*University of Singapore*, [2]*Institute for Infocomm Research, Singapore*.
- **GAO (T1)**: Z. Gao, J. Zhang, L. Zhou, L. Wang, *University of Wollongong, Australia*.
- **LIU (T2)**: L. Liu, S. Paisitkriangkrai, C. Shen, A. van den Hengel, *University of Adelaide, Australia*.
- **MANIVANNAN (T1 and T2)**: S. Manivannan, W. Li, S. Akbar, R. Wang, J. Zhang, S. J. McKenna, *University of Dundee, UK*.
- **NANNI (T1)**: L. Nanni[1], M. Paci[2,3], F.L.C. dos Santos[2,3], J. Hyttinen[2,3], [1]*University of Padua, Italy*, [2]*Tampere University of Technology, Finland*, [3]*BioMediTech, Finland*.
- **PAISITKRIANGKRAI (T1 and T2)**: S. Paisitkriangkrai, L. Liu, R. Hill, C. Shen, A. van den Hengel, *University of Adelaide, Australia*.
- **PONOMAREV (T1 and T2)**: G.V. Ponomarev, M.D. Kazanov, *Russian Academy of Science, Russia*.
- **ROBERTS (T1)**: T. Roberts, *University of Cambridge, UK*.
- **SANSONE (T1 and T2)**: D. Gragnaniello, C. Sansone, L. Verdoliva, *University Federico II of Naples, Italy*.
- **TAORMINA (T1 and T2)**: V. Taormina[1], D. Cascio[2], M. Cipolla[2], F. Fauci[2], G. Raso[2], S. Vasile[1], [1]*CyclopusCAD srl, Palermo, Italy*, [2]*University of Palermo, Italy*.

[2]A simple foreground segmentation such as Otsu's thresholding approach could be used to generate the binary mask from the DAPI image.

- **THEODORAKOPOULOS (T1)**: I. Theodorakopoulos, D. Kastaniotis, *University of Patras, Greece*.

V. Experimental results

The adopted experimental protocol was the following. Each participant received the training set for the both Task 1 and Task 2. The participants used the training set to tune their HEp-2 cell/specimen classification system and then they released the executable(s) for the independent evaluation on the test set. Finally, we ran all the submitted executables on the private test sets collecting the results that are reported forward.

A. Task 1

In Figure 3 we plot the MCA at cell level attained by each method on both the training and the test sets, while in Tables I-XI, we report the confusion matrices of all the methods on the test set.

B. Task 2

In Figure 4 we plot the cells recognition accuracy attained by each method on both the training and the test sets, while in Tables XII-XVIII, we report the confusion matrices of all the methods on the test set.

References

[1] P. Foggia, G. Percannella, P. Soda, and M. Vento. Benchmarking hep-2 cells classification methods. *IEEE transactions on medical imaging*, 2013.

[2] A. S. Wiik, M. Hier-Madsen, J. Forslid, P. Charles, and J. Meyrowitsch. Antinuclear antibodies: A contemporary nomenclature using HEp-2 cells. *Journal of Autoimmunity*, In Press, Corrected Proof.

[3] A. Wiliem, Y. Wong, C. Sanderson, P. Hobson, S. Chen, and B. C. Lovell. Classification of human epithelial type 2 cell indirect immunofluoresence images via codebook based descriptors. In *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013.

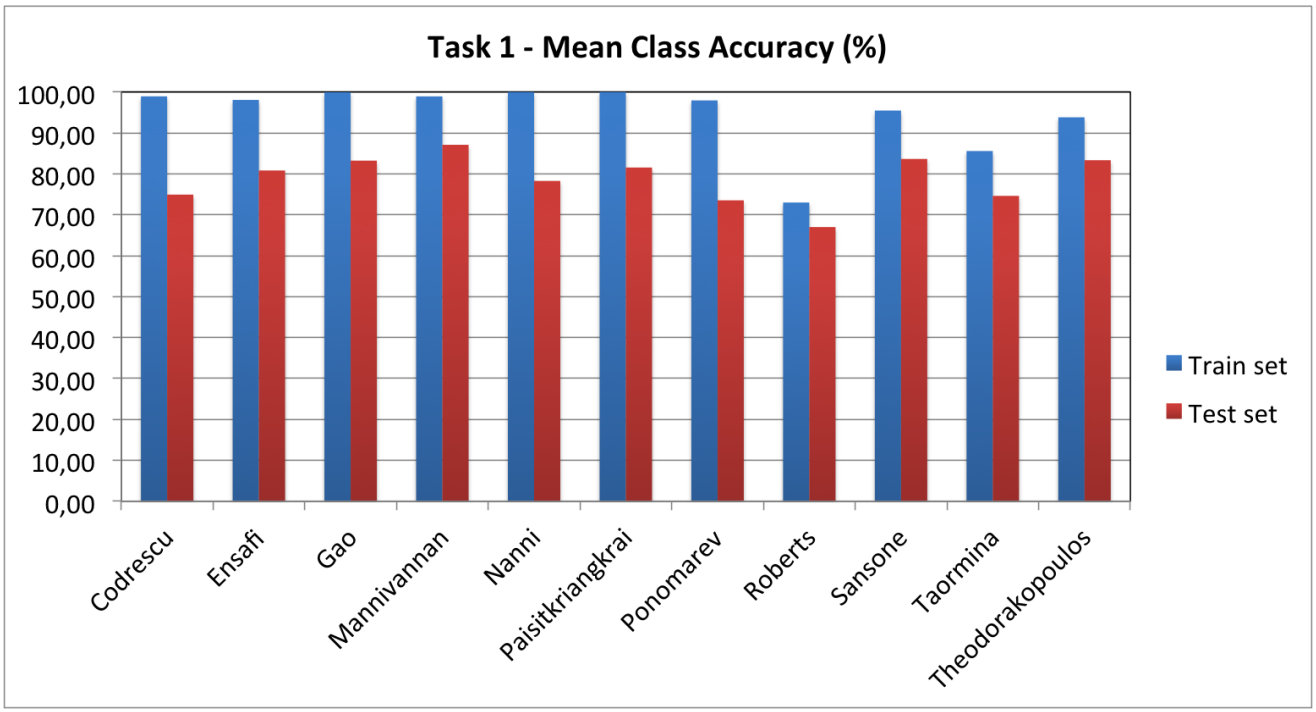


Fig. 3. The MCA at cell level attained by each method on both the training and the test sets.

TABLE I. CONFUSION MATRIX FOR THE CODRESCU'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	91.63	0.64	0.07	3.80	0.66	3.19
Golgi	1.02	55.11	11.84	16.74	12.76	2.53
Homogeneous	0.93	0.43	67.72	3.72	7.45	19.73
Nucleolar	4.11	0.91	1.29	86.46	1.53	5.69
NuMem	0.19	0.88	6.40	2.98	83.69	5.86
Speckled	14.13	0.37	13.42	4.53	2.59	64.98

TABLE II. CONFUSION MATRIX FOR THE ENSAFI'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	94.44	0.69	0.90	1.91	0.64	1.40
Golgi	0.36	70.80	5.13	9.04	12.76	1.91
Homogeneous	0.38	0.72	74.04	5.00	7.84	12.03
Nucleolar	2.51	1.85	1.57	90.18	2.01	1.87
NuMem	0.14	0.60	7.25	0.61	89.41	1.98
Speckled	16.54	0.95	11.33	2.84	2.30	66.04

TABLE III. CONFUSION MATRIX FOR THE GAO'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	96.03	0.18	0.05	1.50	0.48	1.76
Golgi	0.03	73.20	5.75	10.42	9.14	1.45
Homogeneous	0.19	0.84	78.29	5.97	7.52	7.20
Nucleolar	0.72	1.33	1.86	93.72	1.17	1.22
NuMem	0.08	0.83	4.22	0.73	91.27	2.87
Speckled	11.31	0.59	14.61	4.80	1.85	66.85

TABLE IV. CONFUSION MATRIX FOR THE MANIVANNAN'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	97.46	0.13	0.51	0.75	0.22	0.93
Golgi	0.13	81.95	5.39	3.58	8.22	0.72
Homogeneous	0.24	0.79	82.59	5.72	4.52	6.14
Nucleolar	0.75	0.51	1.42	94.75	1.32	1.26
NuMem	0.11	0.64	4.93	0.74	92.19	1.38
Speckled	10.34	0.50	11.99	2.00	1.54	73.64

TABLE V. CONFUSION MATRIX FOR THE NANNI'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	94.72	0.16	0.10	2.40	0.22	2.41
Golgi	0.39	50.61	9.24	14.14	24.93	0.69
Homogeneous	0.50	1.24	69.89	5.87	13.40	9.09
Nucleolar	0.88	0.37	1.62	91.94	3.55	1.63
NuMem	0.28	2.01	2.79	2.36	91.04	1.52
Speckled	10.83	0.44	11.67	2.62	2.99	71.44

TABLE VI. CONFUSION MATRIX FOR THE PAISITKRIANGKRAI'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	94.96	0.12	0.18	1.63	0.61	2.51
Golgi	0.92	65.11	4.21	15.13	11.97	2.66
Homogeneous	0.13	0.21	76.83	6.61	7.56	8.67
Nucleolar	1.02	0.51	1.17	92.92	2.56	1.82
NuMem	0.15	0.65	4.89	1.42	91.08	1.80
Speckled	13.59	0.23	11.91	3.25	2.60	68.42

TABLE VII. CONFUSION MATRIX FOR THE PONOMAREV'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	93.29	0.54	0.17	2.18	0.62	3.19
Golgi	0.95	51.56	6.18	9.93	28.05	3.32
Homogeneous	0.64	2.33	63.34	4.24	17.03	12.42
Nucleolar	2.30	2.70	1.57	86.09	4.97	2.38
NuMem	0.40	2.84	5.41	2.32	85.59	3.45
Speckled	15.03	2.01	12.52	3.49	5.64	61.31

TABLE VIII. CONFUSION MATRIX FOR THE ROBERTS'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	90.70	1.41	0.53	4.21	0.39	2.75
Golgi	0.85	49.49	8.25	11.48	25.49	4.44
Homogeneous	0.99	3.10	58.53	6.16	15.48	15.73
Nucleolar	2.84	7.21	2.15	72.89	7.63	7.29
NuMem	1.17	10.47	7.03	2.86	76.50	1.98
Speckled	11.30	2.71	15.56	11.74	4.83	53.87

TABLE IX. CONFUSION MATRIX FOR THE SANSONE'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	95.52	0.42	0.21	1.15	0.05	2.66
Golgi	0.03	71.82	4.74	7.27	14.60	1.55
Homogeneous	0.05	0.80	78.57	4.94	8.07	7.58
Nucleolar	0.75	1.58	1.96	92.55	1.70	1.46
NuMem	0.05	0.76	3.14	0.85	93.39	1.81
Speckled	13.35	0.71	11.11	2.65	2.17	70.01

TABLE X. CONFUSION MATRIX FOR THE TAORMINA'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	91.81	0.05	0.74	2.03	0.15	5.22
Golgi	0.07	52.58	2.43	14.27	30.25	0.39
Homogeneous	0.02	0.62	73.45	1.74	14.23	9.93
Nucleolar	2.14	2.03	2.92	85.82	3.03	4.06
NuMem	0.28	2.71	8.97	4.57	79.64	3.83
Speckled	15.50	0.40	12.35	2.36	4.97	64.42

TABLE XI. CONFUSION MATRIX FOR THE THEODORAKOPOULOS'S ALGORITHM FOR TASK 1.

%	Centromere	Golgi	Homogeneous	Nucleolar	NuMem	Speckled
Centromere	94.74	0.25	1.31	1.68	0.15	1.87
Golgi	0.30	71.03	5.03	7.53	15.65	0.46
Homogeneous	0.00	0.98	74.31	3.36	13.21	8.14
Nucleolar	0.84	0.92	1.60	92.85	2.24	1.54
NuMem	0.17	1.46	3.64	1.11	91.99	1.63
Speckled	8.18	0.59	12.16	1.69	2.30	75.08

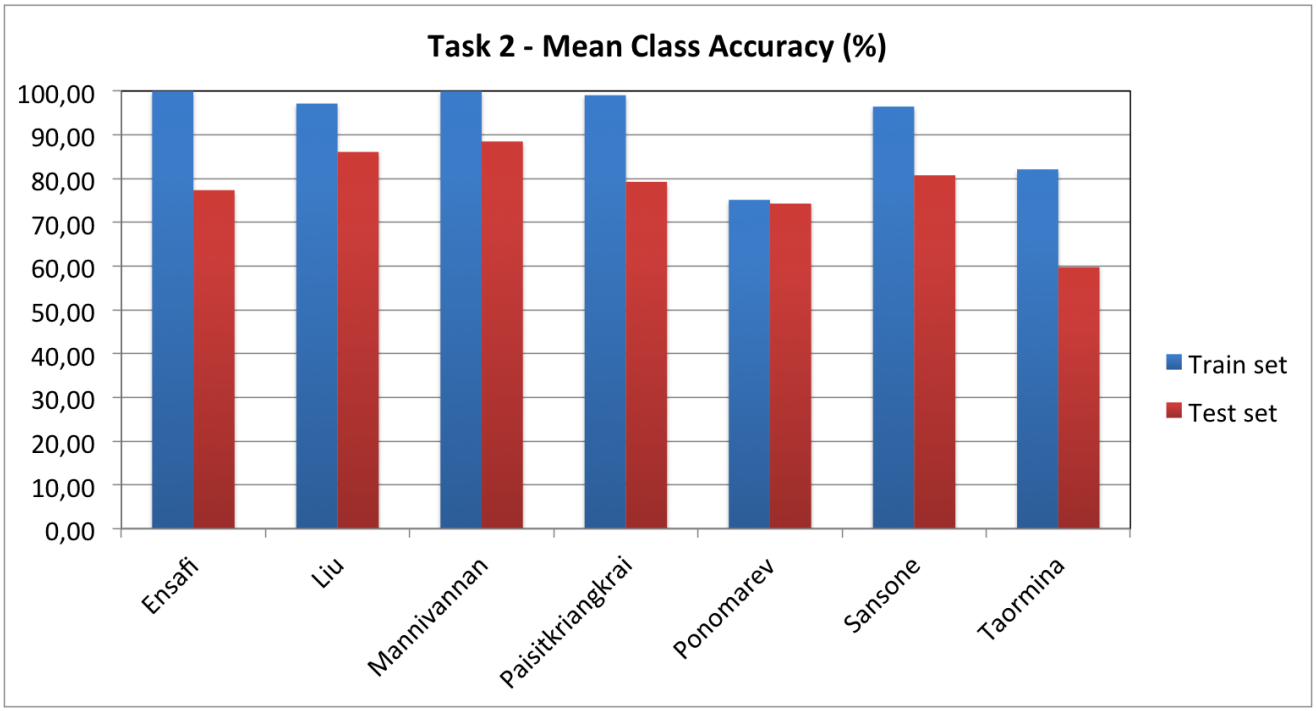


Fig. 4. The MCA at specimen level attained by each method on both the training and the test sets.

TABLE XII. CONFUSION MATRIX FOR THE ENSAFI'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	96.05	0.00	1.32	0.00	1.32	0.00	1.32
Golgi	0.00	73.08	3.85	19.23	0.00	0.00	3.85
Homogeneous	0.00	0.00	82.80	2.55	0.00	1.91	12.74
MitSp	2.27	6.82	40.91	25.00	2.27	15.91	6.82
Nucleolar	0.00	0.00	0.00	0.00	99.33	0.67	0.00
NuMem	0.00	0.00	12.90	0.00	0.00	85.48	1.61
Speckled	4.43	0.00	15.82	0.00	0.00	0.00	79.75

TABLE XIII. CONFUSION MATRIX FOR THE LIU'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	98.68	0.66	0.00	0.00	0.00	0.00	0.66
Golgi	0.00	80.77	7.69	3.85	3.85	3.85	0.00
Homogeneous	0.00	0.00	92.99	0.00	0.00	1.27	5.73
MitSp	0.00	2.27	34.09	52.27	2.27	6.82	2.27
Nucleolar	0.00	0.00	0.67	0.67	98.00	0.67	0.00
NuMem	0.00	0.00	6.45	0.00	0.00	91.94	1.61
Speckled	0.00	0.00	11.39	0.63	0.00	0.00	87.97

TABLE XIV. CONFUSION MATRIX FOR THE MANIVANNAN'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	98.68	0.00	0.00	0.66	0.00	0.00	0.66
Golgi	0.00	80.77	0.00	3.85	3.85	11.54	0.00
Homogeneous	0.00	0.00	92.99	0.00	0.00	1.91	5.10
MitSp	0.00	0.00	18.18	61.36	2.27	11.36	6.82
Nucleolar	0.00	1.33	0.67	0.67	96.00	0.67	0.67
NuMem	0.00	0.00	1.61	0.00	0.00	98.39	0.00
Speckled	0.00	0.00	6.96	0.63	0.00	1.27	91.14

TABLE XV. CONFUSION MATRIX FOR THE PAISITKRIANGKRAI'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	94.74	0.00	0.66	0.00	2.63	0.00	1.97
Golgi	0.00	57.69	0.00	26.92	7.69	3.85	3.85
Homogeneous	0.00	0.00	87.26	1.91	0.00	1.91	8.92
MitSp	2.27	0.00	31.82	50.00	4.55	6.82	4.55
Nucleolar	0.67	0.00	0.00	0.00	96.67	1.33	1.33
NuMem	0.00	0.00	8.06	0.00	0.00	88.71	3.23
Speckled	5.70	0.00	13.92	0.00	0.63	0.00	79.75

TABLE XVI. CONFUSION MATRIX FOR THE PONOMAREV'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	96.05	0.00	0.00	0.00	0.00	0.00	3.95
Golgi	0.00	80.77	0.00	0.00	0.00	19.23	0.00
Homogeneous	0.00	0.00	74.52	0.00	0.00	5.73	19.75
MitSp	0.00	0.00	34.09	0.00	4.55	59.09	2.27
Nucleolar	0.00	1.33	0.00	0.00	94.67	2.00	2.00
NuMem	0.00	0.00	3.23	0.00	0.00	96.77	0.00
Speckled	8.86	0.00	12.03	0.00	0.00	1.90	77.22

TABLE XVII. CONFUSION MATRIX FOR THE SANSONE'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	93.42	0.66	1.32	1.32	1.97	0.66	0.66
Golgi	0.00	73.08	0.00	15.38	0.00	11.54	0.00
Homogeneous	0.00	0.64	74.52	12.10	0.00	3.18	9.55
MitSp	0.00	0.00	13.64	72.73	2.27	6.82	4.55
Nucleolar	1.33	3.33	0.00	4.00	89.33	0.67	1.33
NuMem	0.00	0.00	3.23	4.84	0.00	91.94	0.00
Speckled	5.70	0.00	20.25	2.53	1.27	0.00	70.25

TABLE XVIII. CONFUSION MATRIX FOR THE TAORMINA'S ALGORITHM FOR TASK 2.

%	Centromere	Golgi	Homogeneous	MitSp	Nucleolar	NuMem	Speckled
Centromere	94.08	0.00	0.00	0.00	1.97	0.00	3.95
Golgi	0.00	38.46	15.38	19.23	11.54	11.54	3.85
Homogeneous	0.00	0.00	87.26	0.64	0.00	0.64	11.46
MitSp	0.00	4.55	38.64	20.45	4.55	25.00	6.82
Nucleolar	4.00	1.33	1.33	0.67	80.00	2.67	10.00
NuMem	0.00	0.00	41.94	4.84	12.90	38.71	1.61
Speckled	9.49	0.00	26.58	0.00	4.43	0.00	59.49

Table of content

Quadratic Recurrent FIR MLP for HEP-2 Cells Classification

Cristinel Codrescu
University of Salzburg, Salzburg, Austria, ccodres@cosy.sbg.ac.at

I. Quadratic Recurrent FILMLP

The finite impulse response multilayer perceptron (FIRMLP), a class of temporal processing neural networks, is a multilayer perceptron where the static weights have been replaced by finite impulse response filters. We extended this architecture by replacing the neuron's transfer function from sigmoid to quadratic and by adding a recurrent connection as well. The quadratic transfer function lets the finite impulse response filters act as kernels. The new hybrid temporal architecture consists of layers with different computational properties. First, we introduce the vectorial notation for the network elements. We assume that the synapses in all layers have the same number of taps M^l. Considering a synaptic

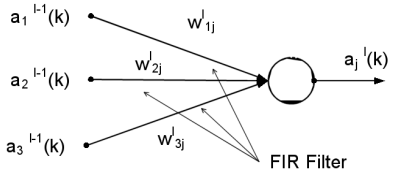


Fig. 1. FIR Neuron vectorial representation

filter that connects neuron *i* to neuron *j* in the layer *l* we denote his coefficients by $w_{ij}^l = [w_{ij}^l(0), w_{ij}^l(1), \dots w_{ij}^l(M^l)]$ and the bias weight with w_{bj}^l. The output of the *j-th* neuron in layer *l* at time *k* is given by the activation vector

$$a_j^l(k) = [a_j^l(k), a_j^l(k-1), \dots a_j^l(k - M^{l+1}) \qquad (1)$$

consisting of the last M^l delayed activation values. By denoting the external inputs as $x_i(k) \widehat{=} a_i^0(k)$ and $y_i(k) \widehat{=} a_i^L(k)$ the network outputs, we can then formulate the reccurent neuron potential in terms of vector dot product in a unified way as

$$s_j^l(k) = \sum_i w_{ij}^l \cdot a_i^{l-1}(k) + w_{bj}^l - wr_j \cdot a_j^l(k-1) \qquad (2)$$

with wr_j the strength of the reccurent connection of the the neuron j.

II. Experiments Description

A. Preprocessing

The original images are not size normalized, the most of them are between 60x60 and 80x80 pixel and represent a rectanglar area not a square. First, we have divided the images into a training and a test set. By randomly choosing 10% from each cell class we have created the test set containing 1359 images. The rest represents the training set. To increase the number of available images we have used two strategies. In the first one, we have scaled the images from the original size to 32x32 pixel. In the second one, we got first the minimum size from the horizontal and vertical dimension of the image. Using this value we have extracted then a centered square from the image and we have scaled it again to 32x32 pixel. By combining the operations flip/rotate we have obtained from each image other seven additional images. Our database contains now 2*8*13,596 = 217,536 images. Each pixel of the grey-level images have been normalized into the range [-1,1].

B. Intermediate-Positive Cell Experiments

In the first experiment set we didn't discriminate between the intermediate and positive cells. For the second we divided the images into four sets. Training and test sets for the intermediate and positive cells. The databases have been created as described before as well. We trained different neural networks as described in II-C for each of the both categories.

C. Neural Networks Setting

The static images contain no temporal information. We can transform this problem in a temporal one by presenting the images to the neural network input in a sliding window procedure. For the experiments we created four layer neural networks. We variate the type of the two first layers (sigmoid, quadratic, reccurent), the number of neurons in each layer, the local connection area size and the neural network error function (mean squared error or cross-entropy). As training algorithm we used online temporal backpropagation adapted to the new architecture.

D. Classification

For the classification task we have created different ensembles of neural networks. The selected networks in each ensemble have been selected by their network error function. We have mean squared error, cross-entropy and mixed (both types) ensemble of networks. For each neural network we have evaluated each test image only once. That means that no flip/rotate operations have been applied on images. After computing the neural network output for a given input pattern, the output neuron with the maximum activation was set to 1 and all others to 0.

References

[1] Badea, C.L. Chapter 3 - FIR Neuron Modeling. In *Ionescu, F. and Stefanoiu, D. (eds.) Approximate dynamic programming for real-time control and neural modeling.*, Stuttgart-Berlin, Germany: Steinbeis Edition (2004)

[2] C. Codrescu. Temporal processing applied to speech recognition. In *Intelligent Systems Design and Applications (ISDA2004)*, Budapest, August 2004.

HEp-2 Cell Image Classification

Shahab Ensafi
School of Electrical and Computer Engineering
National University of Singapore
shahab.ensafi@nus.edu.sg

Shijian Lu
Institute for Infocomm Research
(A*STAR/I2R) Singapore
slu@i2r.a-star.edu.sg

Ashraf Kassim
School of Electrical and Computer Engineering
National University of Singapore
ashraf@nus.edu.sg

Chew Lim Tan
School of Computing
National University of Singapore
tancl@comp.nus.edu.sg

***Abstract*— In this work we present an automatic HEp-2 cell image classification technique that exploits different spatial scaled image representation and sparse coding of SIFT and SURF features. Additionally, spatial max pooling of sparse coding at different scales is used to boost the classification performance. The proposed method is applied on the I3A workshop dataset. Experiments show that it clearly outperforms state-of-the-art techniques in cell and image level as well as two intensity level images.**

Keywords*— *sparse coding; cell image classification

I. METHOD

A. *Task 1*

The framework of the proposed method consists of two stages: training and testing. As only training images are accessible, the cross-validation method is used for testing. In the training stage, the dictionary is learned by using the grid-SIFT and SURF features which are samples from all the training images. Sparse coding is then applied to learn the dictionary and sparse codes iteratively. In the next stage to generate the feature vectors, max pooling is performed on the histogram of the local descriptors in three different scales. Finally a multiclass linear Support Vector Machine (SVM) is learned for image classification. In the testing stage, the same protocol is performed. The sparse codes are obtained by using the learned dictionary and the classification is done by using the trained SVM model.

To improve the accuracy of the classification, the information of the intensity level of images is used as a prior knowledge. In this regard, two different dictionaries are learned by using the *intermediate* and *positive* cell images in the dictionary learning process. The corresponding dictionary is used according to the intensity level of the input image during the test stage.

B. *Task 2*

The method in task 2 is the same as in task 1 whereas the way that the cells are selected to participate in image classification is different. In this task, the cells that the masks are separated from each other are selected and saved in different folder from the 4 images of one specimen. In this stage we have the cells of one specimen and by using the learned dictionaries in previous stage, we can classify the cells. Finally, the maximum vote of all the cells in one specimen is assigned to it as label.

II. RESULTS

A. *Task 1*

The classification accuracy on all the training set is achieved 98.07%. Accordingly, the average accuracy of 97.23% is achieved by performing 7-folded cross-validation method on training set. The results are captured by provided toolbox, which uses Mean Class Accuracy (MCA) and shown in Table 1.

Table 1. The MCA accuracy on training images Task 1.

Class	Accuracy
centromere	99.20%
golgi	100.00%
homogeneous	95.91%
nucleolar	99.50%
numem	99.64%
speckled	94.17%
MCA= 98.07%	

B. *Task 2*

The same evaluation method for task 1 is applied in this task. The accuracy for all training images is achieved 100%. Moreover, the 7-fold cross-validation results in 100% accuracy. The result of MCA is shown in Table 2.

Table 2. The MCA accuracy on training images Task 2.

Class	Accuracy
centromere	100%
golgi	100%
homogeneous	100%
nucleolar	100%
numem	100%
speckled	100%
MCA= 100%	

HEp-2 Cell Image Classification with Convolutional Neural Networks

Zhimin Gao, Jianjia Zhang, Luping Zhou and Lei Wang
School of Computer Science and Software Engineering, Uinversity of Wollongong, NSW 2522, Australia
{zg126, jz163@uowmail.edu.au}, {lupingz, leiw@uow.edu.au}

I. Introduction

Our proposed solution for the HEp-2 cell image classification is based on the deep Convolutional Neural Networks (CNNs)[1], which can automatically extract features from the raw cell image pixels in a hierarchical manner and avoid extracting the classical hand-crafted features. The internal representations for each kind of cell stain patterns are obtained after training the proposed multi-layer networks with provided cell images and class label information via back-propagation[2]. The learnt classification layer in the network predicts the probabilities of a cell image belonging to each of the six classes.

II. Image Preprocessing

Before training our CNNs, two steps of image preprocessing are adopted for each cell image.

Image normalization Each cell image is normalized by first subtracting the minimum intensity value of the image, and then dividing each pixel's intensity by the difference between the nth maximum intensity and the minimum intensity of the image. The normalized intensity values that are larger than 1 are set to 1. We resize the normalized image to the size of 78 x 78 pixels using bicubic interpolation.

Image rotation To increase the number of training samples and the invariance of the obtained networks to cell rotation, we adopt image rotation as our data augmentation[3] method. Each cell image is rotated with an angle step of 18 degree. Hence, the training set is enlarged by a factor of 20.

III. Network Architecture

The architecture of our CNNs contains eight layers, that is, three convolutional layers interlaced with three max-pooling layers, followed with two fully-conneted layers, the last one of which outputs the predictions for an input image. The first layer of the CNNs filters the input image with 6 different kernels of size 7 x 7 with a stride of 1 pixel. A bias is added and a hyperbolic tangent function $f(x) = 1.7159\tanh(\frac{2}{3}x)$ is applied to each unit. They are also applied to the other two convolutional layers and the first fully-connected layer. The second layer takes the output of the first convolutional layer as input, and applies max-pooling over non-overlapping regions of size 2 x 2 for each feature map with a stride of 1. The following convolutional layers and pooling layers also adopt the same stride value of 1 by default. The third layer has 16 kernels of size 4 x 4, and each output feature map is the result of a sum of convolving previous layer's maps with corresponding kernels. The forth layer applies max-pooling over non-overlapping pooling regions of size 3 x 3. The fifth layer includes 32 kernels of size 3 x 3 and operates the same convolution as in the third layer. The sixth layer employs 3 x 3 non-overlapping max-pooling to the output maps of the fifth layer. The resulting 32 feature maps of size 3 x 3 are cascaded and fed into the fully-connected layer (the seventh layer) including 150 neurons. The last layer contains 6 neurons corresponding to each of the six classes, and a softmax activation function is used to produce the predicted probabilities over the 6 classes for each input image.

IV. Network Training

The network is trained on the preprocessed training images. Each pixel of an input image is standardized to have zero mean and unit variance. The weights of the convolutional layers and fully-connected layers are initialized from a uniform distribution and the biases are initialized to zero. All hyper-parameters of the network, such as the number and size of kernels, are chosen from a validation set. All the trainable parameters are updated by stochastic gradient descent to minimize the corss-entropy between the network output and ground truth labels. We train our model starting with a learning rate of 0.01, in conjunction with a momentum of 0.9 to speed up learning, and weight decay of 0.0005 for regularization term. We reduce the learning rate via dividing it by 10 when the training error rate is stabilized. Training is stopped after 100 epochs.

V. Classification

During classification, the test image is also applied the same preprocessing of normalization and rotation as described in section II. For each test image, there will be 20 images totally after rotation. We forward-propagate each of the 20 images through four chosen CNNs, corresponding to networks trained at epoch 75, 85, 95 and 100. The predicted classification label is obtained from the class having maximum probability by first averaging the probabilities of the four CNNs and then averaging the probabilities of the 20 images over the six classes.

References

[1] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.

[2] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel, "Handwritten digit recognition with a back-propagation network," *Advances in Neural Information Processing Systems(NIPS)*, 1990.

[3] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks,", *Advances in Neural Information Processing Systems(NIPS)*, vol. 1, no. 2, 2012.

Specimen Level HEp-2 Image Classification through Majority Voting on Sub-Region Classification Results

Lingqiao Liu, Sakrapee Paisitkriangkrai, Chunhua Shen, Anton van den Hengel

***Abstract*—In our submitted program, we propose to perform the specimen level HEp-2 image classification by using a majority voting framework. More specifically, for each image, we densely extract overlapped subregions and build a classifier to classify them. The final image classification decision is aggregated by performing majority voting on the classification result of each sub-region. The motivation of using sub-region instead of whole image level classification is that in this way we can obtain more training samples and thus it is possible to build a more reliable classifier.**

For each subregion, we densely extract local features by SURF descriptor and use VLAD encoding to obtain the region level representation from the extracted local features. A linear SVM is applied as the classifier. Note that the proposed method does not extract features from every cell but only use all the pixels located in the "cell region". The reason of employing this strategy are twofolds: (1) cell segmentation can be inaccurate when multiple cells are connected. (2) with this strategy, we can better leverage the information from multiple cells.

I. Method Description

A. Overall Framework

The overall structure of the submitted program is depicted in Figure 1. For each image, we densely extract 200 $\times$ 200 sub-regions with the stride of 50 pixels. These sub-regions are then feed into the classifier. The image category is estimated via the majority voting on the decision of each sub-region classifier.

The rational of this strategy is that we only have very limited number of specimen-level images and the classifier trained on them can be very unreliable. By cropping the subregions, we could obtain a large number of samples and much better classifier can be obtained. On the other hand, we do not choose too small sub-region or even performing classifier on the cell level. This is because in this way the classifier decision can be based on multiple cells within a neighborhood. We believe that this scheme is more robust than applying the classifier on the cell level.

B. Local Feature Extraction

For each subregion, we use SURF descriptor [1] and VLAD encoding [2] to obtain the region-level representation. More specifically, we densely extract 8$\times$8 SURF feature with the step size of 4 pixels. The sample points located in the non-cell region (as indicated by the mask image) are discard since they belong to background image.

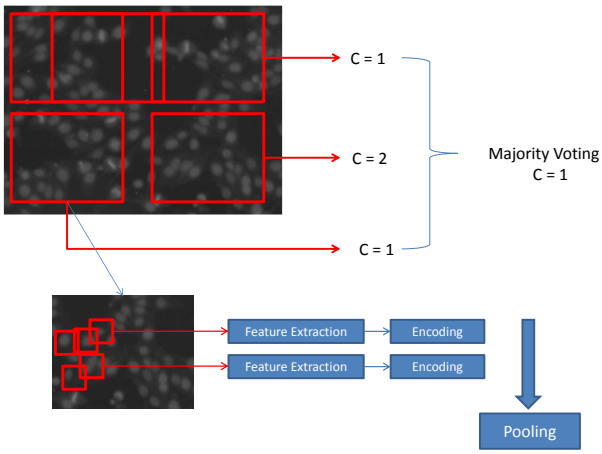


Fig. 1. The illustration of our system.

C. Encoding, Pooling and Classifier Learning

VLAD encoding [2] is employed to encode the extracted SURF feature. In our system, we use k-means algorithm generate a codebook with 100 visual words. Sum-pooling and intra-normalization strategy are used to pool the coding vectors into a vecterized sub-region level representation. Besides that, power normalization is also applied to the pooled coding vector. After obtain the sub-region level feature representation, we train a linear SVM as the region-level classifier.

II. Reference

References

[1] H. Bay, T. Tuytelaars, and L. J. V. Gool. Surf: Speeded up robust features. In *ECCV (1)*, pages 404–417, 2006.

[2] H. Jégou, F. Perronnin, M. Douze, J. Sánchez, P. Pérez, and C. Schmid. Aggregating local image descriptors into compact codes. In *CVPR*, Sept. 2012.

HEp-2 Cell Classification using Multi-resolution Local Patterns and Ensemble SVMs

Siyamalan Manivannan, Wenqi Li, Shazia Akbar, Ruixuan Wang, Jianguo Zhang and Stephen J. McKenna
CVIP, School of Computing, University of Dundee, UK.

***Abstract*—A pattern recognition system was developed to classify immunofluorescence images of HEp-2 cells into six classes: homogeneous, speckled, nucleolar, centromere, golgi, and nuclear membrane. An ensemble of linear support vector machines was used to classify each image. Two-level cell pyramids were used to capture spatial structure and local linear coding was used to encode multiple local features. Leave-one-specimen-out experiments on the I3A Contest *Task 1* data set predicted a mean class accuracy of 79.9%.**

I. Preprocessing and Feature Extraction

We describe a system to classify pre-segmented immunofluorescence images of HEp-2 cells into six classes: homogeneous, speckled, nucleolar, centromere, golgi, and nuclear membrane. Prior to feature extraction, each image was intensity-normalised; specifically, a cell's segmentation mask was dilated (using a 5×5 structuring element) and image intensity values within the dilated mask region were then linearly rescaled so that 2% of pixels became saturated at low and high intensities. Four types of *local* features were then extracted:

- *Multi-resolution local patterns*: a multi-resolution version of local higher-order statistical patterns (LHS). LHS is a non-binarized version of local binary patterns proposed by Sharma et al. [1] for texture classification. It operates on a small image neighbourhood of size 3×3. To capture information from a larger neighbourhood and reduce noise effects, we used the sampling patterns described by Maenpaa [2]; at each sampling point a Gaussian filter was applied, integrating information from the filter's region of support. We call the combination of LHS and these sampling patterns *multi-resolution local patterns*.
- *root-SIFT*: a variant of the widely used SIFT descriptor that has produced better performance than SIFT on some image matching and retrieval tasks [3].
- *Random projections*: projections of patch intensity vectors from the original patch-vector space to a compressed space using randomly chosen projection vectors. Such as scheme has been successfully applied to image texture classification [4]. We used random projections to reduce the dimensionality, d, of a linearised patch to 300 whenever $d > 300$.
- *Intensity histograms*: a $256-$bin intensity histogram was computed from each image patch.

From each image, all features were densely extracted from patches of size 12×12, 16×16, and 20×20 pixels with a step-size of 2 pixels.

A separate $1500-$word dictionary was learned for each feature type using K-means with $300,000$ randomly selected instances of each local feature. Local linear coding, an efficient variant of sparse coding, was used for feature encoding. Max-pooling was used to aggregate the local linear codes. For each feature type, a 2-level cell pyramid was used to capture spatial structure [5]. At the first level, sparse codes from the whole image were pooled to get a feature vector of size 1500. At the second level, feature vectors were computed from the inner region as well as border region of each cell using the segmentation mask information. These three feature vectors were concatenated to give a $4500-$vector. Finally, the four feature types were concatenated to give an $18,000-$vector on which classification was based.

II. Classification

An ensemble of one-vs-rest, multi-class, linear support vector machines was used to classify each cell image. The ensemble consisted of four SVMs, one trained on the original training set images, and others trained on images after they were rotated through $90°$, $180°$, and $270°$ respectively. At test time, each test image was rotated by $0°$, $90°$, $180°$, and $270°$, and each rotated image was then given to the ensemble. This resulted in a set of 16 classification scores for each class. Scores were treated as probabilities using Platt rescaling. The final classification decision was made by averaging these rescaled scores.

Two training regimes were used to generate submissions to the $I3A$ Contest associated with the *ICPR 2014* Workshop. The first regime used only data made available in the Task 1 data set; leave-one-specimen-out experiments with this regime predicted a mean class accuracy of 79.9%. The second regime used an additional set of about 5000 cell images automatically extracted from the Task 2 data set.

References

[1] G. Sharma, S. Hussain, and F. Jurie, "Local higher-order statistics for texture categorization and facial analysis," in *European Conference on Computer Vision*, 2012.
[2] T. Maenpaa, "The local binary pattern approach to texture analysis - extensions and applications," Ph.D. dissertation, University of Oulu, 2003.
[3] R. Arandjelović and A. Zisserman, "Three things everyone should know to improve object retrieval," in *Computer Vision and Pattern Recognition*, 2012.
[4] L. Liu and P. Fieguth, "Texture classification from random features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.
[5] A. Wiliem, B. C. Lovell, P. Hobson, S. Chen, C. Sanderson, and Y. Wong, "Classification of human epithelial type 2 cell indirect immunofluoresence images via codebook based descriptors," in *Workshop on Applications of Computer Vision*, 2013.

HEp-2 Specimen Classification using Multi-resolution Local Patterns and SVM

Siyamalan Manivannan, Wenqi Li, Shazia Akbar, Ruixuan Wang, Jianguo Zhang and Stephen J. McKenna
CVIP, School of Computing, University of Dundee, UK.

***Abstract*—A pattern recognition system was developed to classify immunofluorescence images of HEp-2 specimens into seven classes: homogeneous, speckled, nucleolar, centromere, golgi, nuclear membrane and mitotic spindle. Root-SIFT features together with multi-resolution local patterns were used to capture local shape and texture information. Sparse coding with max-pooling was applied to get an image representation from these local features. Specimens were classified using a linear support vector machine. Leave-one-specimen-out experiments on the I3A Contest *Task 2* data set predicted a mean class accuracy of 87.9%.**

I. Preprocessing and Feature Extraction

We describe a system to classify pre-segmented immunofluorescence images of HEp-2 specimens into seven classes: homogeneous, speckled, nucleolar, centromere, golgi, nuclear membrane and mitotic spindle. Prior to feature extraction, each image was intensity-normalised; specifically, cell masks were dilated (using a 5×5 structuring element) and image intensity values were linearly rescaled so that 2% of pixels within the dilated mask regions became saturated at low and high intensities. Two types of *local* feature were then extracted:

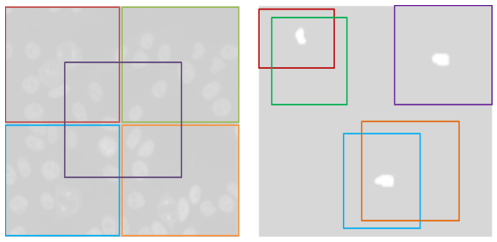
- *Multi-resolution local patterns*: a multi-resolution version of local higher-order statistical patterns (LHS). LHS is a non-binarized version of local binary patterns proposed by Sharma et al. [1] for texture classification. It operates on a small image neighbourhood of size 3×3. To capture information from a larger neighbourhood and reduce noise effects, we used the sampling patterns described by Maenpaa [2]; at each sampling point a Gaussian filter was applied, integrating information from the filter's region of support. We call the combination of LHS and these sampling patterns *multi-resolution local patterns*.
- *root-SIFT*: a variant of the widely used SIFT descriptor that has produced better performance than SIFT on some image matching and retrieval tasks [3].

From each image, all features were densely extracted from image patches using a patch step-size of 4 pixels. Both small patches (12×12 pixels and 16×16 pixels) and large patches (48×48 pixels and 64×64 pixels) were used. Intuitively, small patches can capture local properties at cellular level while large patches can capture information about groups of cells. Features from outside the dilated cell masks were discarded.

A separate $5,000$-word dictionary was learned for each feature type for each patch-size group (i.e., one for *small* and one for *large* patches) using K-means++ with randomly selected instances of each local feature. Local linear coding, an efficient variant of sparse coding, was used for feature encoding. Max-pooling was used to aggregate the local linear codes. The sparse coding feature vectors obtained using the small patch dictionary and the large patch dictionary for each feature type were concatenated to give a $20,000$-vector on which classification was based.

II. Classification

The I3A Contest *Task 2* training set of 1000 images was augmented by including each image after 90° rotation, resulting in a set of 2000 images. Five sub-images were extracted from each image based on the layout shown in Fig 1(a) with the exception of images in the mitotic spindle class. In mitotic spindle images, metaphase cells in which stained mitotic spindle was apparent were manually identified. Five subimages were then extracted around those cells, with some random variation, as shown in Fig 1(b). Finally, the $10,000$ extracted sub-images were added to the $2,000$ images, resulting in an augmented dataset of $12,000$ images for training.



(a) (b)

Fig. 1. Sub-images extracted from specimen images for (a) the homogeneous, speckled, nucleolar, centromere, golgi, and nuclear membrane classes, and (b) the mitotic spindle class.

A one-vs-rest, multi-class SVM classifier was trained on the augmented dataset. At test time, four images for each specimen are given. We run the trained SVM on each of these four images, resulting in four sets of classification scores per specimen. Scores were treated as probabilities using Platt rescaling. The final classification decision was made by averaging the rescaled scores. Leave-one-specimen-out experiments predicted a mean class accuracy of 87.9%.

References

[1] G. Sharma, S. Hussain, and F. Jurie, "Local higher-order statistics for texture categorization and facial analysis," in *ECCV*, 2012.
[2] T. Maenpaa, "The local binary pattern approach to texture analysis - extensions and applications," Ph.D. dissertation, University of Oulu, 2003.
[3] R. Arandjelović and A. Zisserman, "Three things everyone should know to improve object retrieval," in *CVPR*, 2012.

Morphological and Texture Features for HEp-2 Cells Classification

Loris Nanni
loris.nanni@unipd.it
DEI, University of Padua
Padua, Italy

Michelangelo Paci, Florentino Luciano Caetano dos Santos, Jari Hyttinen
ELT, Tampere University of Technology; BioMediTech
Tampere, Finland

***Abstract*—This paper describes our application, aimed to compete for the Cell Level classification task (Task 1). We combined 5 different texture descriptors, extracting from each image 2643 features. Images were classified using support vector machines.**

I. METHOD

Our system for indirect immunofluorescence image classification is based on the combination of 4 descriptors based on Local Binary Pattern (LBP) and 1 morphological feature set: the multiscale Pyramid LBP (PLBP), Local Configuration Pattern (LCP), Rotation Invariant Co-occurrence among adjacent LBP (RIC-LBP), Extended LBP (ELBP) and Strandmark morphological features (STR). The 5 feature sets are classified using Support Vector Machines (SVMs).

A. Preprocessing

Each image in the HEp-2 training set was segmented by its binary mask before the feature extraction.

B. Feature extraction

PLBP [1] is a LBP-derived descriptor based on the Gaussian Pyramid Decomposition of the original image. We considered 3 levels (0, i.e. the original image, 1 and 2) with a 5 by 5 lowpass kernel and a downsampling ratio Rx=Ry=2. $LBP^{riu2}(1,8)$ and $LBP^{riu2}(2,16)$ were considered. *This feature set is made of 84 features.*

STR [2] is a set of morphological (e.g. area, convex Hull area, eccentricity, …) and intensity-based (e.g. mean, median, …) features computed on a collection of images derived from the original image. The collection includes (i) thresholded versions, (ii) cut outs and (iii) magnitude of the gradient of the original image. *This feature set is made of 933 features.*

LCP [3] extracts 2 levels of information: (i) local structural information by LBP and (ii) microscopic configuration (MiC) information. MiC is computed by estimating the optimal weights to linearly reconstruct the central pixel intensity, by exploiting the intensities of the neighboring pixels, minimizing the reconstruction error. Rotation Invariance is achieved by the Fourier transform. *This feature set is made of 370 features.*

RIC-LBP [4] takes into account the spatial relations, i.e. the co-occurrence, among LBP and it is rotation invariant for rotation angles multiple of 45°.We computed RIC-LBP with 3 configuration (LBP radius, displacement among LBPs): (1,2), (2,4) and (4,8). *This feature set contains 408 features.*

ELBP [5] considers both intensity-based and difference-based descriptors. Intensity-based descriptors exploit the central pixel intensity (CI-LBP) or the intensities in the neighborhood (NI-LBP). As for the difference-based descriptors, we only consider the one which exploits the radial distance (RD-LBP). We considered 2 neighborhoods (radius, pixels): (1,8) and (2,16). *This feature set contains 848 features.*

C. Classification

Classification is performed with radial basis function SVMs, according to the "one versus all" approach. Each feature set is classified by 6 SVMs and the 5 partial score corresponding to the same class are combined by the sum rule in one global score, the greatest of which is used to classify the test image. The proposed ensemble outperforms all the tested stand-alone approaches.

REFERENCES

[1] Qian X, Hua X-S, Chen P, Ke L. PLBP: An effective local binary patterns texture descriptor with pyramid representation. Pattern Recognition. 2011;44(10–11):2502–2515.

[2] Strandmark P, Ulen J, Kahl F. HEp-2 staining pattern classification. Pattern Recognition (ICPR), 2012 21st International Conference on. 2012. p. 33–36.

[3] Guo Y, Zhao G, Pietikäinen M. Texture Classification using a Linear Configuration Model based Descriptor. Procedings of the British Machine Vision Conference 2011. British Machine Vision Association; 2011;119.1–119.10.

[4] Nosaka R, Fukui K. HEp-2 cell classification using rotation invariant co-occurrence among local binary patterns. Pattern Recognition. Elsevier; 2014;47(7):2428–2436.

[5] Liu L, Zhao L, Long Y, Kuang G, Fieguth P. Extended local binary patterns for texture classification. Image and Vision Computing. Elsevier B.V.; 2012;30(2):86–99.

Effective HEp-2 Cells Classification using Multi-class Boosting with Heterogeneous features

Sakrapee Paisitkriangkrai, Lingqiao Liu, Rhys Hill, Chunhua Shen and Anton van den Hengel
The Australian Centre for Visual Technologies
The University of Adelaide

***Abstract*—Indirect ImmunoFluorescence (IIF) on human epithelial cells type 2 (HEp-2 cells) is considered as the gold standard for diagnostic screening the presence of autoimmune disease. Unfortunately, this process is known to be labour extensive and time consuming. Computer Aided Diagnosis (CAD) systems plays an important role in addressing these issues. We propose an effective and efficient classification framework to automatically recognize different patterns of HEp-2 cells. We combine a set of diverse and complementary low-level visual features, known to be effective for cell classifications, with a robust and scalable multi-class boosting. The multi-class classifier selects the most discriminative feature at each boosting iteration and combine them into an effective strong classifier. Experimental results show that the method is highly effective and efficient for the task of cell level classification and specimen level classification.**

I. Method description

A. Rationale of the method

In this paper, we describe an effective and efficient classification framework to automatically recognize different patterns of HEp-2 cells. The basic intuition behind our approach is that, instead of using a single feature to discriminate each class from all other classes, it is better to combine a set of diverse and complementary features.

B. Image pre-processing

De-noising and normalization We first apply an image contrast enhancement technique known as histogram equalization.
Image rotation To increase the number of training samples and Improve the robustness of the classifier, each cell image is rotated at every $\pi/4$ degrees angle. For each training image, there will be 7 additional images corresponding to $\{\pi/4, \pi/2, \cdots, 7\pi/4\}$ orientations. During testing, a voting scheme is applied as a post-processing step to increase the robustness of decision making.

C. Feature extraction

Region covariance We propose to use the covariance of several image statistics as the visual descriptor [1]. In this paper, we use the following image statistics: the intensity value; the first and second order derivative in the vertical and horizontal directions; and the magnitude of the gradients.
Local Binary Pattern Due to the large variation in illumination and shape of cells, we use LBP features to capture the characteristics of cells. LBP describes a local region with the magnitude relations between the centre and neighbouring pixel intensities in the local region. In this paper, we use the extension of LBP, known as CoALBP [2], which can describe complex textures by observing not only each local LBP but also the spatial relations among adjacent LBP.
Multi-codebook features Here we assume that local patches of HEp-2 images can be described by a set of linear projections and learn multi-codebook based visual descriptors using HEp-2 cell images. We partition patches into several groups, build multiple bag-of-features models for each group and combine them into the final feature representation. Interested readers should see [3].
Statistical features We adopt the statistical data as described in [4]. The features consist of number of objects in the thresholded image; area; area of convex hull; eccentricity; Euler number and perimeter; mean; standard deviation; entropy; range value; contrast; correlation energy; and homogeneity. Interested readers should see [4] for more details.
Type of Intensities The type of intensities can be one of the following: positive, negative or intermediate.

D. Classification

To achieve a high classification accuracy and real-time performance, we employ the multi-class boosting algorithm of [5] that can adaptively select the most discriminative feature in each boosting iteration and combine them into an effective strong classifier.

E. Submission

Cell level classification We extract heterogeneous features based on the approach described above. Each cell image is classified based on their confidence score.
specimen level classification We apply a scanning window approach to a given HEp-2 specimen image, which contains a variable number of cells. The given image is classified based on a majority voting scheme.

References

[1] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," in *Proc. Eur. Conf. Comp. Vis.*, 2006.
[2] R.Nosaka, Y.Ohkawa, and K.Fukui, "Feature extraction based on co-occurrence of adjacent local binary patterns," in *PSIVT*, 2011.
[3] L. Liu and L. Wang, "Hep-2 cell image classification with multiple linear descriptors," *Pattern Recogn.*, vol. 47, no. 7, pp. 2400–2408, 2014.
[4] P. Strandmark, J. Ulen, and F. Kahl, "HEp-2 staining pattern classification," in *ICPR*, 2012.
[5] S. Paisitkriangkrai, C. Shen, and A. van den Hengel, "A scalable stagewise approach to large-margin multiclass loss-based boosting," *IEEE Trans. Neural Networks.*, vol. 25, no. 5, pp. 1002–1013, 2014.

Image segmentation for ANA HEp-2 classification

Gennady V. Ponomarev*, Marat D. Kazanov*
*Institute for Information Transmission Problems, Russian Academy of Science, Email: krkdl@mail.ru, mkazanov@gmail.com

I. Task 1: ANA HEp-2 cell image classification

The proposed method uses the thresholding and region growing segmentation techniques for segmentation of stained objects and subsequent extraction of object features for image classification (details are in [1]). Several bugs were fixed and the performance of the method was improved.

II. Task 2: specimen image classification

To increase a statistical power of the specimen level image classification we developed an additional preprocessing step for cell separation. For each group of overlapped cells (Figure 1a) we performed: (i) a search for junction points (Figure 1b); (ii) a construction of the separation lines by connecting proper junction points (Figure 1c). A decision for classification of the entire specimen image can be made using both the isolated cells and cells separated by the presented algorithm. We applied unweighted voting scheme for a final classification of the specimen image.

A. Junction points

To detect cell junction points we exploited the fact that all these points are located in the concave segment of the contour and that a maximum of concavity of the contour segment is achieving there. For each point of the contour of the overlapped group of cells we do: (i) a drawing of the line segment by choosing its endpoints as N-th pixels after (A point) and before (B point) the current point C along the contour (Figure 2a) (in our method we fixed N=7); (ii) a calculation of the distance CD from the current point C to AB, which reflects the degree of concavity for a current point. Sign of a distance is considered positive if $\angle CDA$ equals to 90° and negative if $\angle CDA$ equals to -90°. Then, we found local maxima of the CD distances along the contour.

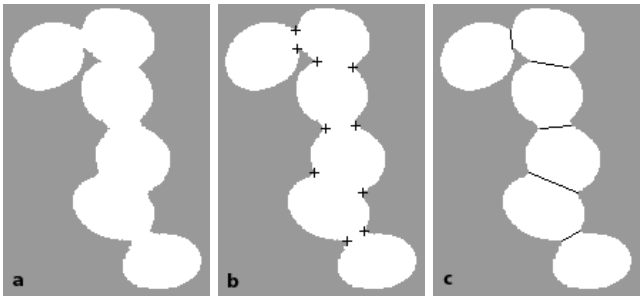


Fig. 1: (a) Image of a group of overlapped cells. (b) Recognized junction points. (c) Final separation lines.

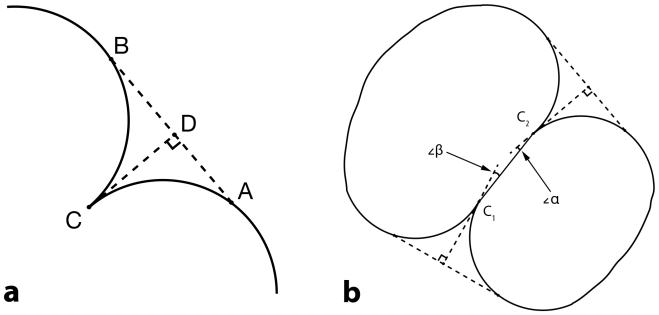


Fig. 2: (a) C is the current pixel of the contour, A/B is the N-th pixel after/before the current pixel along the contour, CD is the normal to AB. (b) Candidate separation line C_1C_2.

Unremarkable maxima, which have been occurred due to noise, were suppressed. Contour points corresponding to the found maxima were assigned as junction points.

B. Separation lines

To find correct connections between junction points for drawing of the separation lines for a particular group of overlapped cells we considered all possible pairs of junction points, a list of pairs L, and estimated for each pair its probability to be a true line, separating one cell from another. For each junction point we already constructed a normal CD to the line segment AB (Figure 2a). Given the nearly elliptical shapes of the cells the line segment AB should be approximately parallel to the line that connects centers of two cell ellipses. Thus, the normal to AB should be an approximation to the line separating two cells. For each pair of junction points we calculated angles $\angle\alpha$ and $\angle\beta$ between line segment connecting two junction points C_1C_2 and two abovementioned lines, respectively (Figure 2b). If two cells were ideal equal ellipses then these three lines were coincided and the angles $\angle\alpha$ and $\angle\beta$ were both equal to zero. Thus, we treated a sum S, $S = \angle\alpha + \angle\beta$, as the measure of deviation of the considered candidate separation line segment from the ideal separation line. The list L was sorted by the S in ascending order. Then, subsequently: (i) a top line segment was selected as a separation line; (ii) all junction point pairs, where selected points are participated, were removed from the list. This procedure was repeated until the list L is empty.

References

[1] G. V. Ponomarev, V. L. Arlazarov, M. S. Gelfand, and M. D. Kazanov, "{ANA} hep-2 cells image classification using number, size, shape and localization of targeted cell regions," *Pattern Recognition*, vol. 47, no. 7, pp. 2360 – 2366, 2014.

Cell Classification using Scattering Coefficients of the Dual-Tree Complex Wavelet Transform

Timothy Roberts

***Abstract*—In this work, we describe how scattering coeffients of a limited-redundancy, directionally-selective wavelet transform can be used to discriminate textures fore the purposes of HEp-2 cell classification.**

I. Notation

In this document we use the following matrix/vector notations:

- $\mathbf{x}_k \in \mathbb{R}^N$ is the k^{th} training vector;
- y_k is the k^{th} class labelling;
- $\mathbf{s}_k \in \mathbb{R}^P$ is the scattering transform of $\mathbf{x}_k$;
- $\mathbf{w}_c$ is a vector of SVM training weights for class c;
- $\mathbf{W}_J$ is a real matrix representing a forward wavelet transformation with J scales;
- $S(\cdot)$ is a forward scattering transform.

II. Algorithm

Our proposed method assumes that cell type classification can be viewed as a texture discrimination problem [1], [2]. We leverage prior work by Bruna et. al [3] in which pooling of the moduli of complex wavelet transform coefficients create translation and scale invariant descriptors for classification.

We consider applying a scattering transform to each k-indexed image $\mathbf{x}_k \in \mathbb{R}^N$ belonging to class $c \in C$ as in [3]. We use the dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) as our sparsifying transform $\mathbf{W}_J$, for its limited redundancy and directionality.

The scattering operator $S(\cdot)$ is formed by M levels of parallel wavelet transforms, followed by taking the modulus of each of the subband coefficients, followed by another set of wavelet transforms on these moduli, and so on. At each level $m \in \{1..M\}$, the coarsest scale coefficients are aggregated by averaging, and the number of scales J in $\mathbf{W}_J$ is modified to accomodate the decreasing resolution of propagated scattering coefficients from one level m to the next. Thus, the aggregation of the coefficients at each scale and level forms a feature vector, which we denote as the scattering representations $\mathbf{s}_k \in \mathbb{R}^P$ corresponding to our original image set.

$$\mathbf{s}_k = S(\mathbf{x}_k) \tag{1}$$

We train a discriminative support vector machine (SVM) to perform classification. In particular, we use "one-versus-the-rest" classification strategy for multi-class SVM classification on the scattering representations, which solves the following program:

$$\arg\min_{\mathbf{w}_c} \|\mathbf{w}_c\|_1 + L \sum_{k=1}^{K} \max(0, 1 - y_k \mathbf{w}_c^T \mathbf{s}_k)^2 \tag{2}$$

where $y_k \in \{-1, +1\}$ represents a class labeling for one of $|C|$ "one-versus-the-rest" optimizations, in which training vectors $\mathbf{s}_k$ belonging to class c are given the label $+1$, and the rest of the training vectors are given the label -1. Repeated execution of this program produces $|C|$ sets of weights, $\mathbf{w}_c$. The algorithm for solving eq. (2) is detailed in [4].

We find that this sparse-weights formulation of the SVM penalty produced the best results experimentally. To perform classification on a test vector $\mathbf{x}_t$, we simply compute $\text{sign}(\mathbf{w}_c^T \mathbf{x}_t) \ \forall \ c \in C$ and use a simple tie-breaking strategy if there exists multiple classes ($+1$ evaluations) for a given training vector.

III. Preprocessing and Parameter Selection

Each test image in the cell classification challenge was segmented using the binary mask supplied with the training data, and centered on a 128×128 template image of all zeros, which is a convenient size for the wavelet transform domain processing used in our method. Parameter selection was performed by repeated experimentation. The parameter set with the smallest classification error was used in the final evaluation version of this algorithm.

Parameter	Value
J $(m = 1)$	4
M	3
L	1

References

[1] Yu-Len Huang, Yu-Lang Jao, Tsu-Yi Hsieh, and Chia-Wei Chung, "Adaptive Automatic Segmentation of HEp-2 Cells in Indirect Immunofluorescence Images," *2008 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (sutc 2008)*, pp. 418–422, June 2008.

[2] Tsu-Yi Hsieh, Yi-Chu Huang, Chia-Wei Chung, and Yu-Len Huang, "HEp-2 cell classification in indirect immunofluorescence images," *2009 7th International Conference on Information, Communications and Signal Processing (ICICS)*, pp. 1–4, Dec. 2009.

[3] Joan Bruna and Stéphane Mallat, "Invariant scattering convolution networks.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1872–86, Aug. 2013.

[4] Rong-en Fan, Xiang-rui Wang, and Chih-jen Lin, "LIBLINEAR : A Library for Large Linear Classification," vol. 9, no. 2008, pp. 1871–1874, 2012.

CELL CLASSIFICATION BY A ROTATION INVARIANT DENSE LOCAL DESCRIPTOR

Diego Gragnaniello, Carlo Sansone and Luisa Verdoliva

University Federico II of Naples, DIETI, 80125 Naples Italy

1. RATIONALE OF THE METHOD

Local descriptors are powerful tools for the analysis of micropatterns and are able to provide a compact and robust represention of the local content. They have shown a high discriminative power in texture classification. Considering that staining patterns are types of different textures, local descriptors are well suited for the task of cell classification. In particular, we decided to use a dense and robust descriptor, hence computed at every point of the image and invariant to some common affine operations.

Among the several descriptors present in the literature, we found that the one proposed in [1] satisfied these requirements. In particular, it has the merit of avoiding to estimate rotation or scale around points that guarantee reliability, like SIFT. In fact, it relies on a combination of log-polar mapping and spatially-varying smoothing. This allows us to convert rotations and rescalings into mere translations. Invariance can be then simply obtained by computing the modulus of the Fourier transform, which is translation invariant.

2. FEATURE EXTRACTION

The local dense descriptor [1] is based on the following steps: 1) log-polar transformation; 2) multiscale smoothing; 3) computation of directional derivatives; 4) Fourier transform along the angular direction (to obtain rotation invariance).

The log-polar grid around each point of the image is formed by 16 angular directions and 10 rings. In order to avoid aliasing caused by the irregular sampling, it is necessary to remove high frequency components by band-pass filtering the image before extracting features from it. This is done by using gaussian filters with spatially varying standard deviations, proportional to the distance from the center of the log-polar sampling grid.

To compute the descriptor, filtering and sampling operate on four directional gradients, rather than on the original values. This is a common practice when building descriptors to provide invariance to intensity changes. In particular, in [1] directional derivatives are computed with polarization, i.e. separating the negative and positive parts. In order to ensure that image rotation amounts to feature translation, directional derivatives need to be aligned with the ray directions. In addition, since the amplitude of spatial derivatives decreases with scale, due to the scale-space smoothing, the output is rescaled by the kernel's standard deviation. All the measurements on the grid are rearranged in a matrix and a Fourier transform along the angular direction is performed to ensure rotation invariance. By exploting the symmetry of the Fourier transform the features can be halved. Finally, concatenating all components in a single vector will form the local descriptor.

3. CLASSIFICATION

Bag of Words model (BoW) has shown its superiority over many conventional global features for image classification, hence we used it in this context.

We computed the euclidean distance between the feature vector (truly we extracted the feature vectors inside the segmentation map) and the codewords of the dictionary to build the histogram that will be the input of a SVM linear classifier. We considered a different BoW cardinality for each class, and in particular 200 for Centromere and Nuclear membrane, 100 for Nucleolar and Speckled, and 50 for Golgi and Homogeneous. Since soft-assignment can improve performance, as shown in [2], we followed this path by considering gaussian weights.

Using the described approach we were able to obtain a mean class accuracy of 81.76 % on the training set, by performing a leave-one-out per subject cross-validation procedure.

4. REFERENCES

[1] I. Kokkinos, M.M. Bronstein, and A. Yuille, "Dense scale invariant descriptors for images and surface," Tech. Rep., INRIA, 2012.

[2] J.C. van Gemert, J.-M. Geusebroek, C.J. Veenman, and A.W.M. Smeulders, "Kernel codebooks for scene categorization," in *European Conference on Computer Vision*, pp. 696–709. Springer, 2008.

HEp-2 Cell classification using multi-process based on KNN

Taormina Vincenzo[1], Cascio Donato[2], Cipolla Marco[2], Fauci Francesco[2], Raso Giuseppe[2], Vasile Simone Maria[1]

[1] CyclopusCAD srl, Palermo, Italy

[2] Dipartimento di Fisica e Chimica, Università Degli Studi di Palermo, Palermo, Italy

Abstract— We present a feature extraction and classification scheme to classify the fluorescence staining patterns of HEp-2 cells in IIF images. We propose a set of complementary processes specific to each class of patterns to search. We propose an classification-approach based on two steps: the first of the type one-against-all, the second of the type one-against-one. To do this, we needed to implement 21 KNN classifiers: 6 OAA and 15 OAO.

***Keywords*— IIF images, K–Nearest-Neighbors (K-NN), multi-class, classification, one-against-all classification, leave-one out cross validation.**

In several multiclass classification problems, it is preferable to use a number of classifier equal to the number of classes and each classifier is trained to discriminate a class from all the others (binary approach). In this work, in addition to differentiate the classification stage by implementing six classifiers for six classes of staining patterns, the preprocessing and feature extraction steps are differentiated, too. We adopt, then, a non-standard pipeline for supervised image classification. Figure 1 shows the flow of operations adopted in this work. The generic image is simultaneously processed by six processes obtaining six separate outputs that represent how the cell resembles each one of the 6 classes analyzed in this work. The choice of methods, features and parameters was performed automatically, using the Mean Class Accuracy (MCA) as a figure of merit. The parameters involved are tuned using a cross validation scheme. The main benefit of this pipeline is that it has a good explanatory power, while its principle is simple to explain. Different staining patterns can be characterized by a limited set of attributes describing the spatial relationships between pixels values and the main image variations occurring in each cell type: this information is generally obtained by means of textural analysis techniques. These techniques can be grouped into two major categories: (i) statistical methods describing the distribution of grey-levels in the image; and (ii) frequency domain measurements of image variations [1-2]. Reducing the dimensionality of the data by selecting a subset of the original variables may be advantageous for simplifying the classification problem from different points of view.

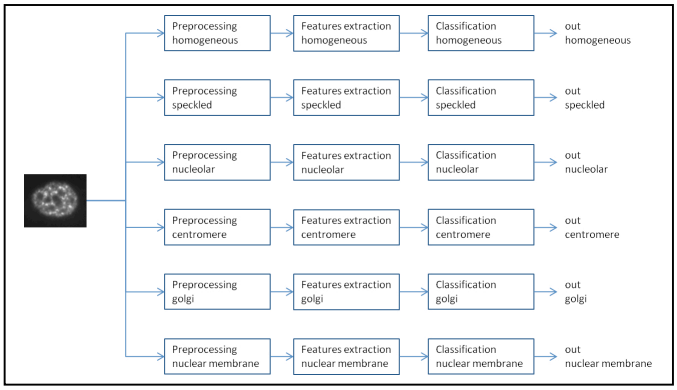


Fig. Pipeline of developed method.

Our choice was to sort the extracted features by using the Fisher values and select only the 15 highest values for each pattern class; each pattern class will be described by a set of 15 features (out of 108) where such set may be different among the classes. We propose, in this paper, an approach based on two steps of classification: the first of the type OAA, the second (and final) of the type OAO. To do this, we needed to implement 21 KNN classifiers: 6 one-against-all and 15 one-against-one. The objective of the second step is to correct any errors that occurred in the first step. For each classifier the training and validation sets are used for the optimization and cross validation phases. The leave-one-out cell cross validation [3] technique is used to exploit the highest possible number of patterns during the training phase, without invalidate the results.

References

[1] Masala, G.L et al. Classifiers trained on dissimilarity representation of medical pattern: A comparative study. *Nuovo Cimento C*, 28, 905-912, 2005.

[2] Masala, G. L. et al. Comparative study of feature classification methods for mass lesion recognition in digitized mammograms. Nuovo Cimento C 2007, 30, 305-316.

[3] Cascio, D., Magro, R., Fauci, F., Iacomi, M., Raso, G. Automatic detection of lung nodules in CT datasets based on stable 3D mass-spring models. *Computers in Biology and Medicine*, 42 (11) , pp. 1098-1109, 2012.

HEP-2 Cells classification in IIF images using K-Nearest Neighbours

Taormina Vincenzo[1], Cascio Donato[2], Cipolla Marco[2], Fauci Francesco[2], Raso Giuseppe[2], Vasile Simone Maria[1]

[1] CyclopusCAD srl, Palermo, Italy

[2] Dipartimento di Fisica e Chimica, Università Degli Studi di Palermo, Palermo, Italy

Abstract— In this paper we presente a structured process of classification based on K-Nearest Neighbors. The basic intuition behind our approach is that, instead of using a single process to discriminate each class from all other classes, it is better to combine a set of different and complementary processes. The method is based on the KNN classifiers.

***Keywords*— IIF images, K–Nearest-Neighbors (K-NN), multi-class, classification, one-against-all classification, leave-one out cross validation.**

In several multiclass classification problems, it is preferable to use a number of classifier equal to the number of classes and each classifier is trained to discriminate a class from all the others (binary approach). In this work, in addition to differentiate the classification stage by implementing six classifiers for six classes of staining patterns, the preprocessing and feature extraction steps are differentiated, too. We adopt, then, a non-standard pipeline for supervised image classification. The generic image is simultaneously processed by six processes obtaining six separate outputs that represent how the cell resembles each one of the 6 classes analyzed in this work. The choice of methods, features and parameters was performed automatically, using the Mean Class Accuracy (MCA) as a figure of merit. The parameters involved are tuned using a cross validation scheme. Different types of preprocessing have been analyzed and for each class it has been identified the preprocessing giving the best performance in terms of final class accuracy. Table 1 shows the details of combinations of preprocessing analyzed and their possible use. The choice of the best preprocessing for a given class of staining patterns has been carried out to maximize the MCA. We combine several discriminative visual features known to be effective for cell classifications with a robust and scalable multi-class boosting. In our work we extract a large number (108) of features [1-2] able to fully characterize the staining pattern of HEp-2 cells. We implement 21 KNN classifiers: 6 one-against-all and 15 one-against-one. The leave-one-out cell cross validation [3] technique is used to exploit the highest possible number of patterns during the training phase, without invalidate the results.

TABLE I. LIST OF PREPROCCESING ANALYZED AND THEIR USE

	First function Filtering	***Second Function Contrast normalization***	***Used for pattern:***
1	None	None	---
2	None	Contrast normalization	homogeneos
3	Dilation 5x5	None	---
4	Dilation 5x5	Contrast normalization	centromere
5	Dilation 9x9	Contrast normalization	nucleolar
6	Erosion 5x5	None	---
7	Erosion 5x5	Contrast normalization	---
8	Erosion 9x9	Contrast normalization	---
9	Median 5x5	None	---
10	Median 5x5	Contrast normalization	---
11	Median 9x9	Contrast normalization	---
12	Gaussian 5x5	None	---
13	Gaussian 5x5	Contrast normalization	golgi, nuclear menbrane
14	Gaussian 9x9	Contrast normalization	speckled

References

[1] Masala, G.L et al. Classifiers trained on dissimilarity representtation of medical pattern: A comparative study. *Nuovo Cimento C*, 28, 905-912, 2005.

[2] Masala, G. L. et al. Comparative study of feature classification methods for mass lesion recognition in digitized mammograms. Nuovo Cimento C 2007, 30, 305-316.

[3] Cascio, D., Magro, R., Fauci, F., Iacomi, M., Raso, G. Automatic detection of lung nodules in CT datasets based on stable 3D mass-spring models. *Computers in Biology and Medicine*, 42 (11) , pp. 1098-1109, 2012.

HEp-2 Cells Classification using morphological and rotation-invariant textural features

Ilias Theodorakopoulos and Dimitris Kastaniotis

Electronics Laboratory, Physics Department, University of Patras, Greece.

{ iltheodorako, dkastaniotis }@upatras.gr

I. RATIONALE OF THE METHOD

Staining patterns presented on IIF slides are characterized by great variability. Fluorescence intensity is distributed either in structured, stochastic or intermediate manner. Every fluorescence pattern group is characterized by unique optical properties originating from the nature of the depicted cells.

In order to efficiently capture these characteristics, a set of morphological features which is a variant of the two dimensional Boolean texture models is incorporated, along with well-known textural descriptors appropriately modified in order to encode textural information in a rotation-invariant manner.

II. IMAGE PREPROCESSING

On each cell image a median filter is applied in order to eliminate isolated intensity extremities. The original image is then normalized by subtracting the minimum value of the filtered image, and then dividing each pixel's value by the difference between maximum and minimum intensity of the filtered image.

Subsequently, a noise-reduction step is applied by dividing the cell image into overlapping patches which are encoded using sparse representation techniques with a dictionary learned on low-noise images, in order to suppress noise. The cell image is subsequently reconstructed, and the pixel values are re-normalized in the range 0 to 1.

III. FEATURE EXTRACTION

A. Mophrological features

A set of binary images are constructed via application of thresholding operation to the image, utilizing a set of 14 equally spaced values in the range of [0,1] as threshold values. Significant amount of information is carried by the resulting sequence of binary images regarding the spatial distribution of intensities on the depicted staining pattern. In order to quantify this information, Connected Component Analysis is performed in each binary image, and the following set of morphological features is computed for each of them: Number of detected objects, density in binary image and mean objects' solidity, where objects of size less than 1% of the mean objects' size of each binary image, are considered as noise and ignored during the calculation of the above features. Finally, the complexity of the cell's contour is considered as an additional feature expressed as the difference between the cell's contour and the perimeter of the equivalent circle. Further details on the incorporated features can be found on [1].

B. Textural Features

Rotation-invariant SIFT descriptors are computed on a dense grind over the cell image, similar to [2]. Then the distribution of the 128-dimensional descriptors is encoded according to the VLAD scheme in order to form a compact descriptor similar to [3].

Additionally, three variants of Local Binary Patterns where incorporated in order to further encode textural information in several scales. Uniform LBPs, Four-Patch LBPS (FPLBPs) and Center-Symmetric LBPs (CS-LBPs) are computed on the cell image, and the corresponding binary patterns are rotated so as the least significant bit of each pattern to be located closest to the principal direction of the local gradient. Finally, histograms of the corresponding LBPs where computed in order to form the descriptors.

IV. CLASSIFICATION

All descriptors computed on a cell image are concatenated into a vector of 1296 dimensions. The resulting feature vector is normalized by subtracting the mean vector and dividing each feature by the standard deviation of the corresponding values from the training set. The final classification is performed using a multiclass SVM scheme with linear kernel.

REFERENCES

[1] I. Theodorakopoulos, D. Kastaniotis, G. Economou, S. Fotopoulos, "HEp-2 Cells classification via fusion of morphological and textural features," Bioinformatics & Bioengineering (BIBE), 2012 IEEE 12th International Conference on , vol., no., pp.689,694, 11-13 Nov. 2012. doi: 10.1109/BIBE.2012.6399750.

[2] I. Theodorakopoulos, D. Kastaniotis, G. Economou, S. Fotopoulos, "HEp-2 cells classification via sparse representation of textural features fused into dissimilarity space," Pattern Recognition, vol. 47, no. 7, pp 2367-2378, July 2014. http://dx.doi.org/10.1016/j.patcog.2013.09.026

[3] D. Kastaniotis, I. Theodorakopoulos, G. Economou, S. Fotopoulos, "HEp-2 cells classification using locally aggregated features mapped in the dissimilarity space," Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on , vol., no., pp.1,4, 10-13 Nov. 2013. doi: 10.1109/BIBE.2013.6701591