Computers in Biology and Medicine 42 (2012) 376–386

Contents lists available at SciVerse ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Real time decision support system for diagnosis of rare cancers, trained in parallel, on a graphics processing unit

Konstantinos Sidiropoulos [a,*], Dimitrios Glotsos [b], Spiros Kostopoulos [b], Panagiota Ravazoula [c], Ioannis Kalatzis [b], Dionisis Cavouras [b], John Stonham [a]

[a] School of Engineering and Design, Brunel University West London, Uxbridge, Middlesex, UB8 3PH, UK
[b] Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos, Egaleo, Athens, 12210, Greece
[c] Department of Pathology, University Hospital of Patras, Rio, Patras 265 00, Greece

ARTICLE INFO

Article history:
Received 8 July 2011
Accepted 2 December 2011

Keywords:
Decision support system
Parallel processing
Graphics processing unit (GPU)
Rare brain cancers

ABSTRACT

In the present study a new strategy is introduced for designing and developing of an efficient dynamic Decision Support System (DSS) for supporting rare cancers decision making. The proposed DSS operates on a Graphics Processing Unit (GPU) and it is capable of adjusting its design in real time based on user-defined clinical questions in contrast to standard CPU implementations that are limited by processing and memory constrains. The core of the proposed DSS was a Probabilistic Neural Network classifier and was evaluated on 140 rare brain cancer cases, regarding its ability to predict tumors' malignancy, using a panel of 20 morphological and textural features Generalization was estimated using an external 10-fold cross-validation. The proposed GPU-based DSS achieved significantly higher training speed, outperforming the CPU-based system by a factor that ranged from 267 to 288 times. System design was optimized using a combination of 4 textural and morphological features with 78.6% overall accuracy, whereas system generalization was 73.8% $\pm$ 3.2%. By exploiting the inherently parallel architecture of a consumer level GPU, the proposed approach enables real time, optimal design of a DSS for any user-defined clinical question for improving diagnostic assessments, prognostic relevance and concordance rates for rare cancers in clinical practice.

1. Introduction

The majority of rare cancers are diagnosed on the basis of histological evaluation. Even though histological diagnosis is fundamentally important for patient management, the potential of diagnostic errors still remains substantially high, ranging from 25% to 40%, in routine conditions[1–3]. Factors affecting diagnostic accuracy include experts' subjectivity and lack of experience, intra and inter observer variability and poor tumor sampling; rare cancers low prevalence and their biological complexity hinder the establishment of concrete criteria able to predict tumors' behavior, and, thus, to administrate proper treatments. The latter might explain the fact that a/ although promising treatments have been proposed [4–6], death rates have not been yet reduced[4,7–9] and b/ the cost of rare cancers management still remains one of the highest healthcare economic burdens in Europe and worldwide [10,11]. There is a pressing need to develop methods for improving diagnostic accuracy in order to reduce diagnostic errors and, thus, guide successful choice of therapies. Accurate diagnosis linked to proper therapeutic strategies would eventually benefit patients with rare cancers improving survival and quality of life, while at the same time, keeping healthcare costs at an acceptable level.

Although it has long been recognized that (a) more resources need to be committed towards rare cancers, considering that even common types of cancer manifest multiple rare subtypes, (b) rare cancers don't conform to strict guidelines, thus experience from similar cases would improve management, (c) reproducibility in histological typing is a significant liability, since absence of well-defined criteria lead to different management decisions based on the subjective interpretation of different experts, only few studies have focused on developing new strategies for improving health care services to patients suffering from rare cancer. Research has suggested that possible solutions can be found in training and education of clinicians [12], telepathology as a result of synergistic consensus effort of several clinicians [13], and decision support systems using quantitative features extracted by image analysis [14–19].

DSS, in particular, have an increasing impact in clinical decision making based on histopathological images. A great deal

* Correspondence to: Medical Image and Signal Processing (medisp) Lab., Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos, Egaleo, Athens, 12210, Greece. Tel.: +30 2105385375.
E-mail addresses: Konstantinos.Sidiropoulos@brunel.ac.uk, ksidis@gmail.com (K. Sidiropoulos).
URLS: http://www.teiath.gr/stef/tio/medisp/index.htm, http://medisp.bme.teiath.gr/ (K. Sidiropoulos).

0010-4825/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2011.12.004

of financial resources has been allocated worldwide towards promoting DSS-based projects, such as the eTumour [20], and the DISHEART [21] projects. Very few projects have been funded for histological and histopathological decision making (such as the EUROPATH [22], E-SCOPE [23] and TUBAFROST [24] projects) and almost no integrated attempt has been made towards exploiting DSS for better health care services for rare cancers. On the other hand, literature shows piecemeal approaches on the basis of DSS focusing on specific subtypes of rare tumors, such as astrocytomas, glioblastomas, meningiomas and atypical metastatic cancers [14–19]. Although such attempts have highlighted the contribution of DSS in improving diagnostic assessments, prognostic relevance and concordance rates, the exploitation of such systems in clinical practice is limited due to the following reasons: First, proposed DSS have been designed to focus on specific groups (two or three class problems) of rare cancers [14–19]. However, the exact classification of each type of rare tumor is, in most cases, a result of an exhaustive search from a list of multiple possible outcomes based on the cell type of origin of the disease (for example the new World Health Organization classification of tumors affecting the central nervous system defining nine categories of neuroepithelial tumors, each of which has subcategories dictating choice of therapy and prognosis [25]). Second, reliable DSS require a considerable number of cases for learning and designing of the prediction rules. However, due to the low prevalence of rare cancers [26], it is difficult to collect large datasets that will ensure reliability in decision making. Third, quantitative descriptor of tumoral microscopic images comprises a large list of structural, morphological, textural and ordinal assessments. The effective selection of the most informative of these descriptors would require an exhaustive search among all possible combinations. However, exhaustive search is limited by serious computational and memory limitations.

In the present study a new strategy is introduced for designing and developing of an efficient dynamic decision support system for supporting rare cancers decision making. The proposed DSS is built on a Graphics Processing Unit (GPU) framework and it is capable of updating its structure in real time whenever a new verified case is uploaded on its repository. Moreover, the system adjusts its design based on the exact clinical question to be answered, which is user-defined from on a list of potential outcomes according to existing clinical classifications. Finally, the GPU-framework enables the optimal selection of the most informative histological typing criteria, in contrast to standard CPU implementations that are limited by processing and memory constrains. In this way, it is possible to examine exhaustively a significant larger panel of quantitative histological descriptors combining structural, textural and morphological information, something that is unrealistic on standard CPU implementations. The proposed DSS was evaluated on rare brain cancer cases.

2. Material and methods

2.1. Material

Clinical material comprised 140 cases of rare brain tumors, including astrocytomas (variants: gemistocytic, fibrillary and mixed), anaplastic astrocytomas, glioblastomas, oligodendroglioma and mixed gliomas. At least three Hematoxylin-Eosin (H&E) stained sections were generated from the same block of formalin-fixed paraffin-embedded tissue for each case (patient). The type and the degree of malignancy (grade) of each tumor was defined according to the WHO guidelines for classification of neuroepithelial tumors of the central nervous system[25]. Of the 140 cases, 61 were characterized as grade II (aggressive), 35 as grade III (more aggressive) and 44 as grade IV (most aggressive). Each grade category consisted of subtypes listed in Table 1.

Five images for each case were digitized based on the guidelines of an experienced histopathologist (P.R.). The digitization was performed using a light microscopy imaging system consisted of a Zeiss Axiostar-Plus microscope (ZEISS; Germany) connected to a Leica DC 300F color video camera (LEICA; Germany).

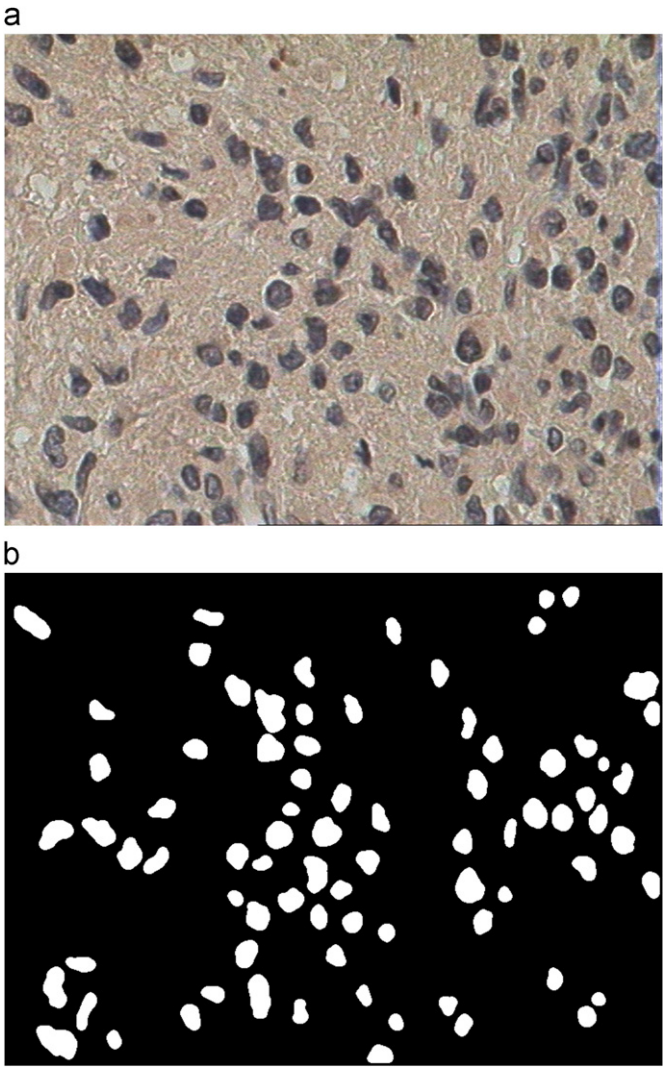

Fig. 1. (a) An example of a high grade pleomorphic glioblastoma multiforme tumor. (b) Segmentation of Figure1a for isolation of nuclei from surrounding tissue.

Table 1
Brain tumors dataset.

Astrocytoma (WHO grade II)			Astrocytoma (WHO grade III)		Glioblastoma multiforme (WHO grade IV)		
Gemistocytic	Fibrillary	Mixed	Anaplastic	Anaplastic mixed	Giant cell	Gliosarcoma	Pleomorphic
8	19	34	25	10	13	1	30
	61			35		44	

2.2. Methods

Images were converted to greyscale for further processing. A pixel-based classification algorithm presented by our group elsewhere [27] was used to isolate nuclei from surrounding tissue. An example of an original and a segmented image of a grade IV pleomorphic glioblastoma multiforme is illustrated at Figs. 1a and b respectively. Segmented nuclei were quantified by means of 20 textural and morphological features describing the size, shape and grain texture of nuclei. The list of the features computed is presented at Fig. 2. The latter features have been shown of diagnostic and prognostic significance in grading and typing of brain tumors[15,18].

The pattern recognition system was designed with a Probabilistic Neural Network (PNN) [28] classifier with Gaussian kernel (the spread of the Gaussian kernel was experimentally determined equal to 0.2), the leave-one-out [29] and the exhaustive search [30] to estimate the resubstitution error and based on this error select the most informative features. An external 10-fold cross validation process was employed for estimating the generalization of the prediction rule to unknown cases. According to the external cross validation data were randomly split into 10 different non-overlapping subsets. At each stage of the cross validation the leave-one-out and the exhaustive search method was implemented to 9 datasets, external to the tenth test subset. The average external cross validation error was calculated over 50 such random splits.

In order to adapt the training of the aforementioned classifier to the inherently parallel SIMD (Single Instruction Multiple Data) architecture of a GPU, the whole procedure was broken into many tasks that were designed to run concurrently. In technical terms, the challenge was to design the kernel, or the small code fragment running in multiple parallel threads, in an optimal way in order to evenly distribute the workload and maximize performance. It should be noted that in the present study NVIDIA's Compute Unified Device Architecture (CUDA) [19] version 4.0 programming framework was selected and utilized, mainly due to its maturity.

In the proposed implementation, the first step involved the enumeration of all possible feature combinations. Following the transfer of the training dataset from the memory of the host PC to the GPU's global memory, each thread was assigned with a single feature combination (Fig. 3). Specifically, the task of each thread, running concurrently, was to train the PNN classifier with this unique feature combination and evaluate its classification accuracy by means of the leave-one-out technique. Hence, for a given training dataset consisting of N patterns, every thread trained the PNN classifier N times for a single combination of features, but each time leaving another pattern out of the training set. It should be noted that feature combinations were transferred to the device memory for processing in batches of 80,000 combinations. Upon completion of all threads, results were transferred back to the host's memory for presentation (Fig. 4). Listing 1 shows the kernel function executed by each CUDA thread. Block and Grid size were experimentally set to $[128 \times 1]$ and $[625 \times 1 \times 1]$ respectively.

In order to evaluate its performance, the developed GPU-based DSS was trained in a series of experiments and the required training time was measured. The latter was then compared with the respective training time of the same classifier running on a typical CPU and programmed in C programming language. Two series of experiments were performed in order to investigate the impact of both a/feature dimensionality and b/dataset size in training time. Hence, regarding the first series of experiments, both systems were trained with a dataset, consisting of 140 patterns that progressively increased their feature dimensions, beginning with 14 features and reaching 20 by a step of 2. In the second series of experiments the two systems were trained with a

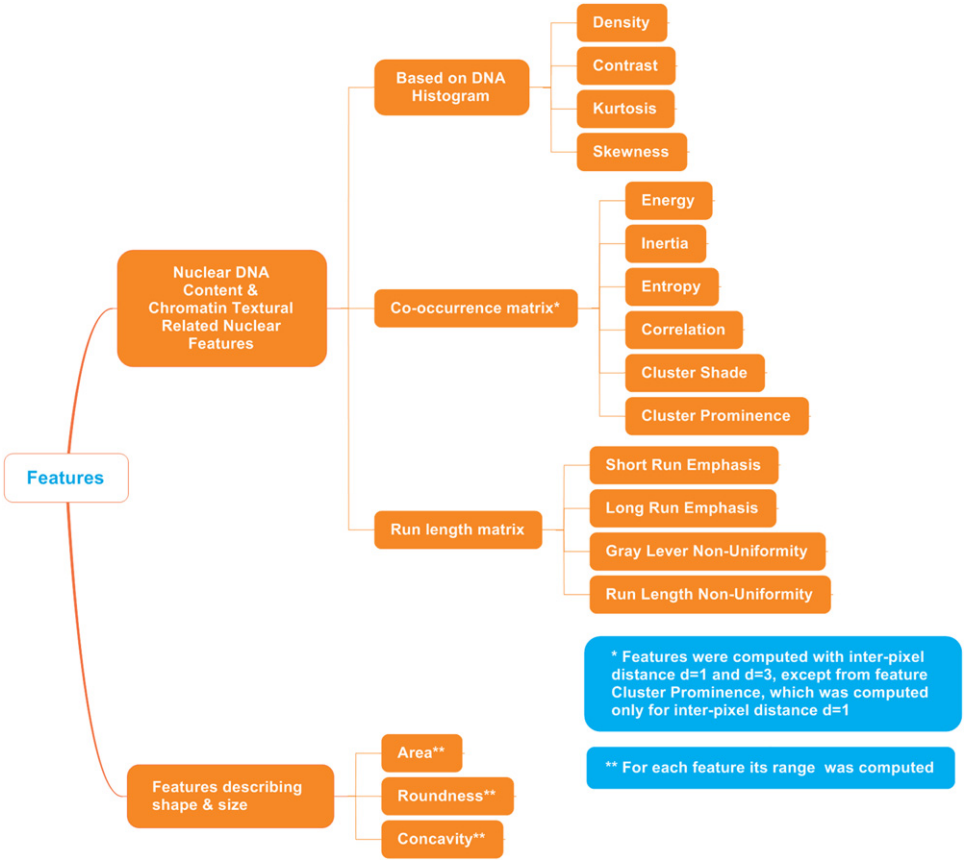

Fig. 2. Morphological, textural and structural tumoral quantitative descriptors.

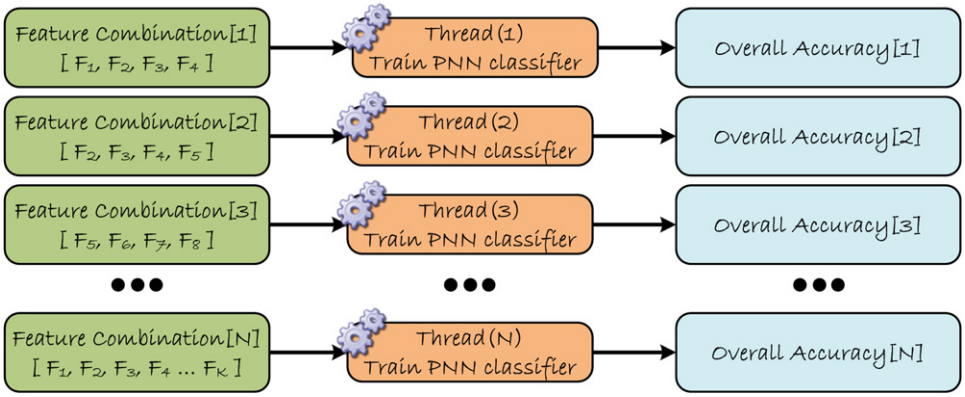

Fig. 3. Parallel training of the proposed GPU-based DSS distributed to N threads, where each GPU thread was assigned with the training of a PNN classifier for a unique feature combination.

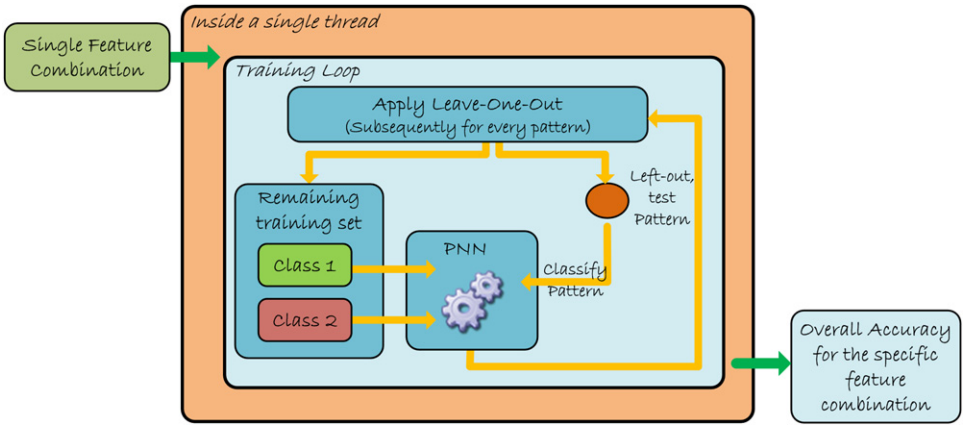

Fig. 4. The task of each thread, running concurrently, was to train the PNN classifier with a unique feature combination and evaluate its classification accuracy by means of the leave-one-pattern-out technique. In every case the output of each thread was the overall accuracy achieved for the specific feature combination.

Table 2
Impact of dataset's feature dimension in designing the PNN classifier using the exhaustive search and the leave-one-out methods.

Total number of features exhaustively combined	Number of distinct feature combinations	PNN Training Time for 140 patterns (ms)				CPU/GPU Training. Time
		CPU	GPU			
			Processing	Transfer	Total	
14	16,383	266,696	994.50	4.01	998.51	267.1
16	65,535	1,202,711	4,164.91	10.88	4,175.79	288.0
18	262,143	5,351,447	18,475.58	46.75	18,522.33	288.9
20	1,048,575	23,557,518	84,421.21	139.76	84,560.97	278.6

20-feature dataset that progressively increased its size, by adding more patterns. Therefore, beginning with 80 patterns, the dataset steadily increased its size to 140 patterns by steps of 10.

All experiments were performed on a desktop PC featuring an Intel Core 2 Quad CPU at 2.83 GHz, and hosting a GeForce GTX 580. The hardware specification of the proposed method experimental setup is illustrated in Tables A1 and A2 (see Appendix A).

3. Results

The first series of experiments attempted to investigate the impact of feature dimensionality in the performance of the proposed DSS. Table 2 illustrates the computation time measured by both the GPU and the CPU-based systems. As far as the GPU is concerned, total processing time includes also the time required for memory transfers between the host's and the GPU's memory. As one can easily observe, the proposed GPU-based PNN classifier system achieved significantly higher training speed in all cases, outperforming the CPU-based system by a factor that ranged from 267.1 to 288.9 times (Fig. 5). In case of exhaustive combination of 20 features, which translates to 1048575 distinct feature combinations, the CPU-based system required about 6.5 h while the GPU-based system completed training in just 85 s.

The purpose of the second series of experiments was to assess the effect of the dataset size on the required training time of the proposed GPU-based DSS. According to the results, presented in Table 3 and illustrated in Fig. 6, training time of both systems is increased at almost the same rates, with the GPU-based DSS outperforming the CPU-based system by a factor that ranged from 278.5 to 279.4 times.

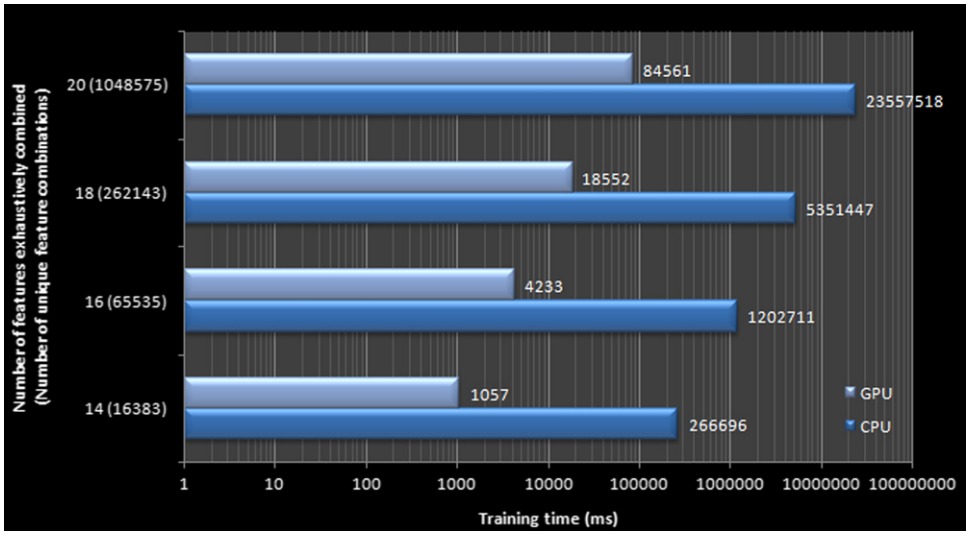

Fig. 5. Comparative assessment of CPU and GPU regarding training time of the DSS versus the dataset's feature dimensionality (the *x*-axis is presented in logarithmic scale).

Table 3
Impact of dataset's size in designing the PNN classifier using the exhaustive search and the leave-one-out methods.

Number of Patterns	PNN Training Time (ms)				CPU/GPU Training. Time
	CPU	GPU			
		Processing	Transfer	Total	
80	7,695,287	27,455.52	138.64	27,594.16	278.9
90	9,756,267	34,786.87	137.40	34,924.27	279.4
100	12,010,644	42,946.45	137.65	43,084.10	278.8
110	14,544,027	52,083.37	138.32	52,221.69	278.5
120	17,306,957	61,887.07	136.92	62,023.99	279.0
130	20,289,976	72,725.09	138.77	72,863.86	278.5
140	23,557,518	84,421.21	139.76	84,560.97	278.6

It should be noted, that in all experiments performed, transfer time varied between 0.17% and 0.50% of total training time depending on the number of features combined and the number of patterns constituting the training set.

According to the exhaustive search and the LOO method, the proposed GPU-based DSS gave optimum results in the discrimination of low from high grade rare brain tumors using 4 features: low grade cases were correctly identified with 72.1%, high grade cases with 83.5% whereas the overall accuracy achieved was 78.6%. The performance of the system for different number of feature combinations is illustrated in Fig. 7. Applying the external 10-fold cross-validation method, overall accuracy was $73.8\% \pm 3.2\%$.

4. Discussion

DSS have been proven as important second opinion tools for reducing inter-observer reliability and diagnostic misinterpretations in diagnostic pathology[31–33]. Considering that accurate histological diagnosis leads to more precise treatment planning in rare cancers, DSS, as well as other tools contributing towards this direction, are of crucial importance in improving chances of survival of patients. Although piecemeal research efforts have been presented in literature regarding DSS for rare cancers [14–19], their application into clinical practice is limited due to the following reasons: (a) available data are limited due to low prevalence of such diseases; small sample sizes are prohibitive for designing reliable DSS. (b) rare cancers recognition is a multiclass problem; it is very difficult to build a DSS system able to identify all possible outcomes of any type of rare cancer, since it would require a significant amount of data for each subtype of tumor whose identification is of clinical importance. (c) quantitative descriptors extracted from biopsy images include a panel of structural, textural, morphological and qualitative expert-based ordinal scales. The optimal selection of the best subset of these descriptors would require exhaustive search algorithms to seek all possible combinations. However, the running time of an exhaustive search is proportional to the dimension and size of the dataset. An exhaustive search within 20 features for 140 data with the leave-one-out method would require the designing of more than 140 million classifiers. In practice common pattern recognition applications employ much less than 20 features in exhaustive search and stop the search at combinations of up to 7–8 due to memory and CPU restrictions.

One of the solutions proposed, so as to tackle the aforementioned problem, is parallel processing typically involving powerful supercomputers, or server clusters. Unfortunately, this kind of hardware is expensive and therefore accessible only to few people. However, a new promising development in this regard is the emergence of consumer-level graphics processing units (GPUs) as a mainstream computing platform[34].

Previous studies in the field of image processing and analysis that attempt to benefit in speed from the utilization of GPUs, include implementations of neural networks [35], Kernel methods for Support Vector Machine classifiers [36], k-Nearest Neighbor search methods [37], and algorithms for both computed tomography reconstruction [34] and registration of medical images [38,39]. In another study [40], Ruiz et al. proposed a GPU-based CAP system for prognosis of neuroblastoma achieving 45 times faster execution time compared to a CPU based system.

The proposed DSS offers solutions to the above problems. First, it is designed using an external cross-validation process to ensure generalization to unseen data [41]. External cross validation is the most plausible candidate method in cases where we don't have the luxury of withholding a part of data for testing. In this way, two goals are achieved: the DSS can be designed with limited data, which is the case in rare cancers, while safeguarding that success rates will apply to new cases. Moreover, the system can be re-designed in real time when new verified cases are added to its repository increasing, in this way, the sample size and the experience of the system in

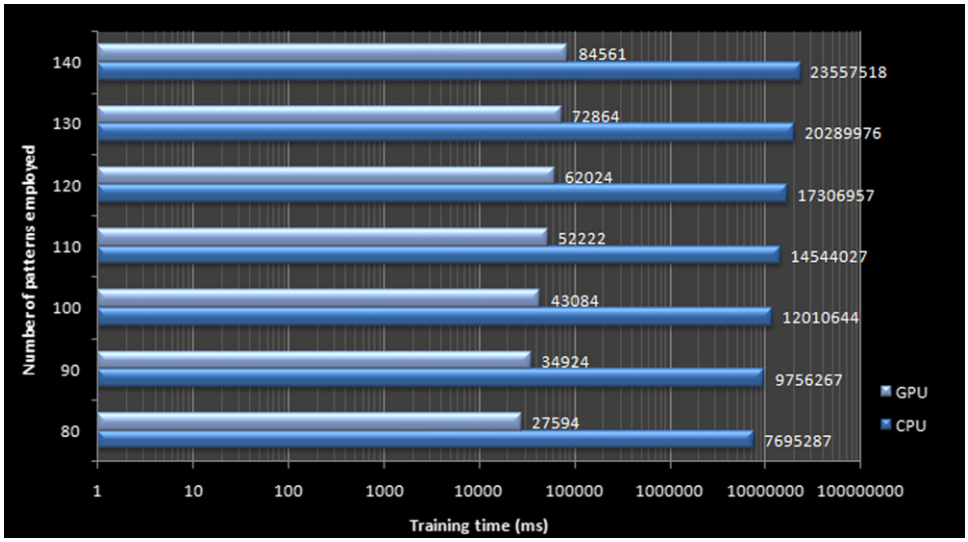

Fig. 6. Comparative assessment of CPU and GPU regarding training time of the DSS versus the dataset's size (the *x*-axis is presented in logarithmic scale).

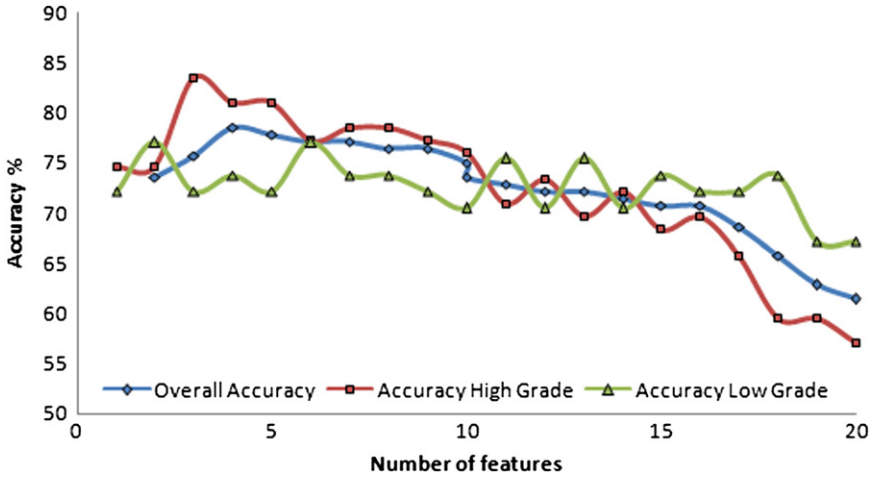

Fig. 7. Success rates of the proposed GPU-based DSS in the discrimination of low from high grade rare brain tumors.

predicting clinical meaningful subcategories. The time needed to design the system for 140 cases on a CPU required 6.5 h, whereas in for the GPU-framework it took around 1.4 min (see Table 2 or Fig. 5). Second, the fast and dynamic implementation on the GPU-framework enables the expert to define the exact categories that need to be separated. In most rare cancers diagnostic difficulties arise from problems of tumoral biological continuum; thus, the degeneration of the multi-class problem to two-class differential diagnosis problem is a reasonable assumption. Under this framework, the expert physician may design in real time a DSS for separation of any subtype of rare cancers using the same dataset. From Table 3 or Fig. 6 it can be observed that the GPU is more than 278 times faster than the CPU, enabling, in this way, real time exploitation of the system. Third, it is possible to investigate the performance of higher dimension datasets on the basis of the optimal exhaustive search method. In this way, the expert may review the predictive value of complementary information from different origin tumoral descriptors, such as textural, morphological, structural and ordinal descriptors. In the experiments for brain tumor grade determination optimum results (78.6%) were obtained using combinations of 4 features (concavity, density, energy and inertia). Higher feature combinations resulted in inferior prediction rates. This result was anticipated since as the number of features increases for fix sample size, data become sparse and the curse of dimensionality predominates. The rule of thumb indicates that the ratio of dataset size to the number of features should not fall under 20 [42]; this was experimentally verified since in our case best results were obtained for the low grade class when the aforementioned ratio was equal to 20.3 (61 cases and 3 features), whereas for the high grade class when the respective ratio was 19.8 (79 cases and 4 features) (see Fig. 7). The latter is a key observation regarding the need to re-train the DSS each time a new batch of cases is added to its repository. Re-training on larger sample sizes would potentially enable the identification of an even larger feature subset of improved performance than those already identified, without being limited by peaking phenomena effects. In this way, the investigation of complementary information of different origin features would be feasible.

To test the generalization of the proposed method to new data, a 10-fold external cross-validation procedure was applied, giving estimates up to $73.8\% \pm 3.2\%$ overall accuracy. The reduced performance of the system under the external cross validation process is expected since the leave-one-out method leads to optimistically biased estimates of the classifier performance. The leave-one-out method can be used for efficient classifier design by means of selecting most informative features and best settings for the prediction rule; however, if one seeks to estimate the generalization capability of the prediction rule to unseen cases, then the external validation rate is more realistic than the leave-one-out rate.

Conflict of interest statement

None declared

Listing 1–The kernel used to design the proposed DSS.

```
__global__ void
GPUSingleCombination_LOO_PNN(
// Pointer to the dataset stored in global memory
float* features,
// Pointer to the mean values lookup table stored in global memory
float* meanFeatures,
// Pointer to the standard deviations lookup table stored in global memory
float* stdvFeatures,
// Size of dataset – Number of patterns
int D_SIZE,
// Defines the point in dataset where class 2 begins
int CDIVIDER,
// Max number of features
int P_SIZE ,
// Pointer to all possible feature combinations stored in global memory
int* combination ,
// Pointer to the size of each feature combination
int* combinationSize,
// Pointer to an array used to store the overall accuracy
// for each possible feature combination
float* overallAccuracy )
{
  //Unique thread identifier
  int index = (blockIdx.x * 128) + threadIdx.x;
  //Various variables
  int i, k, c1;
  unsigned int  combK;
  float unknownP[22];
  int class1Count, class2Count;

  float p = 2.0f;
```

```
float sigma = 0.2f;
float pi = 3.14159f;
float distance, distanceInP;
float denominator = 2.0f * powf(sigma,2);
float denominatorP = 0.0f;
float tempSum, tempExpSum, expvar;
int classSelected=0;
float g1 = 0.0;
float g2 = 0.0;
int truthTable[2][2];
truthTable[0][0] =0;
truthTable[0][1] =0;
truthTable[1][0] =0;
truthTable[1][1] =0;

//Get the size of the specific combination
unsigned int COMBINATION_SIZE = combinationSize[index];

// For all patterns
// Apply Leave one out method
for(i=0; i<D_SIZE; i++){

  classSelected=1;
   if(i<CDIVIDER)
      classSelected=0;
   //If 0 then patterns belongs to Class 1 else it belongs to Class 2
   // Copy unknown pattern from global memory and normalize it
   for(k=0; k<P_SIZE; k++){
      unknownP[k] = ( features[(i * P_SIZE) + k]
               - meanFeatures[(i * P_SIZE) + k] )  / (float) stdvFeatures[(i * P_SIZE) + k];
   }
   // Calculate discriminant PNN function for Class 1
   class1Count = 0;
   tempExpSum = 0.0;
   for(c1=0; c1<CDIVIDER; c1++) {

      if(c1!=i){ // NOT the left-out pattern
```

```
      class1Count++;
      // Calculate square distance
      tempSum = 0.0;
      for(k=0; k<COMBINATION_SIZE; k++){
         combK = combination[ (index * P_SIZE) + k ];
        distance =  unknownP[ combK ] - ( ( features[ (c1 * P_SIZE) + combK ] -
                    meanFeatures[ (i * P_SIZE) + combK  ] ) / (float)
                    stdvFeatures[  (i * P_SIZE) + combK  ]  );
        distanceInP = powf( distance , 2);
        tempSum = tempSum + distanceInP;
      }
      expvar = expf( - tempSum/denominator);
      tempExpSum = tempExpSum + expvar;

   }//if(c1!=i)

}

denominatorP = powf( 2.0f * pi , (COMBINATION_SIZE/2) ) * powf( sigma , COMBINATION_SIZE);
g1 = tempExpSum / ( denominatorP * class1Count);
// Calculate discriminant PNN function for Class 2
class2Count = 0;
tempExpSum = 0.0;
for(c1=CDIVIDER; c1<D_SIZE; c1++) {
   if(c1!=i){ // NOT the left-out pattern
      class2Count++;
      // Calculate square distance
      tempSum = 0.0;
      for(k=0; k<COMBINATION_SIZE; k++){
         combK = combination[ (index * P_SIZE) + k ];
                    distance =  unknownP[ combK ] - ( ( features[ (c1 * P_SIZE) + combK ] -
                    meanFeatures[ (i * P_SIZE) + combK  ] ) / (float)
                    stdvFeatures[  (i * P_SIZE) + combK  ]  );
         distanceInP = powf( distance , 2);
         tempSum = tempSum + distanceInP;
      }
      expvar = expf( - tempSum/denominator);
```

```
            tempExpSum = tempExpSum + expvar;


         }//if(c1!=i)
      }
      denominatorP = powf( 2.0f * pi , (COMBINATION_SIZE/2) ) * powf( sigma , COMBINATION_SIZE);
      g2 = tempExpSum / ( denominatorP * class2Count);
      //Fill up the truth table
      if(g1>=g2){
         truthTable[classSelected][0]++;
      }
      else{
         truthTable[classSelected][1]++;
      }
   }
   //Calculate OverallAccuracy for this feature combination
   float OverallAccuracy = 0.0f;
   OverallAccuracy = (truthTable[0][0] + truthTable[1][1]) * 100
         / (truthTable[0][0] + truthTable[0][1] + truthTable[1][0] + truthTable[1][1] );
   //Update the value in global memory
   overallAccuracy[index] = OverallAccuracy;
   //Wait for all threads to reach this point
   __syncthreads();
}
```

Acknowledgments

The first author was supported by a grant from the Greek State Scholarships Foundation (IKY).

Appendix A

See Tables A1 and A2.

Table A1
Specifications of the host PC.

Host hardware specification	
Processor	Intel Core 2 Quad, 2.83 GHz
Memory	4 GB DDR3
Motherboard	ASUSTeK P5Q3
Operating System	Windows 7 Ultimate 32 bit

Table A2
Specifications of GPU device.

GPU specification	
Model	GeForce GTX 580
Number of multiprocessors	16
Number of CUDA cores	512
Total global memory	1472 MB GDDR5
Shared memory per block	48 KB
Memory interface width	384 bits
Memory clock speed	4200 MHz
Warp size	32
Core clock speed	832 MHz
CUDA compute capability	2.0

References

[1] C. Fletcher, Diagnostic Histopathology of Tumors, Churchill Livingstone, Edinburgh, New York, 1995.
[2] A.D. Ramsay, Errors by locums. Histopathology departments already audit diagnostic errors, Br. Med. J. 313 (1996) 117.
[3] D.B. Troxel, Diagnostic errors in surgical pathology uncovered by a review of malpractice claims. Part III. breast biopsies, Int. J. Surg. Pathol. 8 (2000) 335–337.

[4] F. Cowie, Treatment of rare cancers: gastrointestinal stromal tumours, Br. J. Hosp. Med. 67 (2006) 361–364.
[5] J. Gonzalez, M.R. Gilbert, Treatment of astrocytomas, Curr. Opin. Neurobiol. 18 (2005) 632–638.
[6] W.E. Longo, A.M. Vernava 3rd, T.P. Wade, et al., Rare anal canal cancers in the U.S. veteran: patterns of disease and results of treatment, Am. Surg. 61 (1995) 495–500.
[7] J.A. DiSario, Colorectal cancers of rare histologic types compared with adenocarcinomas, Dis. Colon Rectum 38 (1995) 1227.
[8] A. Dixon, Rare skin cancers in general practice, Aust. Fam. Physician 36 (2007) 141–143.
[9] T. Joannides, Rare cancers, Clin. Oncol. 13 (2001) 235.
[10] Y. Belkacemi, A. Zouhair, M. Ozsahin, D. Azria, R.O. Mirimanoff, [Prognostic factors and management of rare cancers], Cancer/Radiothér. 10 (2006) 323–329.
[11] L. Kutikova, L. Bowman, S. Chang, et al., Utilization and cost of health care services associated with primary malignant brain tumors in the United States, J. Neuro-oncol. 81 (2007) 61–65.
[12] R.A. Prayson, D.P. Agamanolis, M.L. Cohen, et al., Interobserver reproducibility among neuropathologists and surgical pathologists in fibrillary astrocytoma grading, J. Neurol. Sci. 175 (2000) 33–39.
[13] W. Coons, P. Jhonson, B. Sceithauer, A. Yates, D. Pearl, Improving diagnostic accuracy and interobserver concordance in the classification and grading of Primary Gliomas, Cancer 79 (1997) 1381–1393.
[14] N. Belacel, M. Boulassel, Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis, Artif. Intell. Med. 21 (2001) 201–207.
[15] C. Decaestecker, I. Camby, N. Nagy, et al., Improving morphology-based malignancy grading schemes in astrocytic tumors by means of computer-assisted techniques, Brain Pathol. 8 (1998) 29–38.
[16] Guo-Zheng Li, Yang Jie, Ye Chen-Zhou, Geng Dao-Ying, Degree prediction of malignancy in brain glioma using support vector machines, Comput. Biol. Med. 36 (2006) 313–325.
[17] R. Nafe, W. Schlote, B. Schneider, Histomorphometry of tumour cell nuclei in astrocytomas using shape analysis, densitometry and topometric analysis, Neuropathol. Appl. Neurobiol. 31 (2005) 34–44.
[18] P.K. Sallinen, S.L. Sallinen, P.T. Helen, et al., Grading of diffusely infiltrating astrocytomas by quantitative histopathology, cell proliferation and image cytometric DNA analysis, Comparison of 133 tumours in the context of the WHO 1979 and WHO 1993 grading schemes, Neuropathology and Applied Neurobiology 26 (2000) 319–331.
[19] M. Scarpelli, R. Montironi, D. Thompson, P. Bartels, Computer-Assisted Discrimination of Glioblastomas, Anal. Quant. Cytol. Histopathol. 19 (1997) 369–375.
[20] eTumour, 〈http://www.ist-world.org/ProjectDetails.aspx?ProjectId=ec81f1ded2ff4c82a6573cc77f2a3c79〉.
[21] DISHEART, 〈http://www.ist-world.org/ProjectDetails.aspx?ProjectId=2916094e804346ecbffc9d9a45fe9cf8〉.
[22] EUROPATH, 〈http://telescan.nki.nl/action/europath.htm〉.
[23] E-SCOPE, 〈http://cordis.europa.eu/fetch?CALLER=PROJ_ICT&ACTION=D&CAT=PROJ&RCN=71240〉.
[24] TUBAFROST, 〈http://www.tubafrost.org/〉.
[25] A. Rousseau, K. Mokhtari, C. Duyckaerts, The 2007 WHO classification of tumors of the central nervous system – what has changed? Curr. Opin. Neurobiol. 21 (2008) 720–727.
[26] Paolo G. Casali, Do rare cancers deserve specific strategies for cancer research?, The Lancet Oncology 11 506–507.
[27] D. Glotsos, P. Spyridonos, D. Cavouras, et al., Automated segmentation of routinely hematoxylin-eosin-stained microscopic images by combining support vector machine clustering and active contour models, Anal. Quant. Cytol. Histol. 26 (2004) 331–340.
[28] D.F. Specht, Probabilistic, Neural Net. 3 (1990) 109–118.
[29] P.A. Lachenbruch, An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis, Biometrics 23 (1967) 639–645.
[30] J. Fellow, R. Duin, J. Mao, Statistical pattern recognition: a review, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 4–37.
[31] Dorin Comaniciu, Bogdan Georgescu, Peter Meer, Wenjin Chen, David Foran, Decision Support system for multiuser remote microscopy; in: Telepathology 12th IEEE Symposium on Computer-Based Medical Systems (CBMS '99), 18–20 June 1999, IEEE Computer Society, Stamford, CT, USA.
[32] R. Logeswaran, Cholangiocarcinoma–an automated preliminary detection system using MLP, J. Med. Syst. 33 (2009) 413–421.
[33] P. Spyridonos, D. Cavouras, P. Ravazoula, G. Nikiforidis, A computer-based diagnostic and prognostic system for assessing urinary bladder tumour grade and predicting cancer recurrence, Med. Inf. Internet Med. 27 (2002) 111–122.
[34] F. Xu, K. Mueller, Real-time 3D computed tomographic reconstruction using commodity graphics hardware, Phys. Med. Biol. 52 (2007) 3405–3419.
[35] Oh. Kyoung-Su, Keechul Jung, GPU implementation of neural networks, Pattern Recognit. 37 (2004) 1311–1314.
[36] Julius. Ohmer, Maire, Frederic. and Brown, Ross., Implementation of Kernel Methods on the GPU digital imaging computing: techniques and applications (DICTA 2005) 2005.
[37] Vincent Garcia, Eric Debreuve, Ross Brown, Fast k nearest neighbor search using GPU, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24–26 June 2008, IEEE Computer Society, Anchorage, Alaska, USA.
[38] R. Shams, P. Sadeghi, R. Kennedy, R. Hartley, Parallel computation of mutual information on the GPU with application to real-time registration of 3D medical images, Comput. Meth. Prog. Bio. (2009).
[39] R.J. Lapeer, S.K. Shah, R.S. Rowland, An optimised radial basis function algorithm for fast non-rigid registration of medical images, Comput. Biol. Med. 40 (2010) 1–7.
[40] A. Ruiz, O. Sertel, M. Ujaldon, et al., Stroma classification for neuroblastoma on graphics processors, Int.J.Data Mining Bioinf. 3 (2009) 280–298.
[41] C. Ambroise, G.J. McLachlan, Selection bias in gene extraction on the basis of microarray gene-expression data, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 6562–6566.
[42] V. Kechman, Learn. Soft Comput. (2001) 121–184.