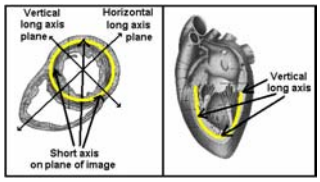
Lukasz Kurgan

Intelligent Systems Laboratory
NC 3015

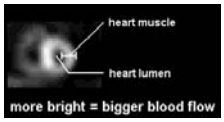
Ensemble of Classifiers to Improve Accuracy of SPECT Heart Image Analysis System

SPECT Heart Image Analysis System

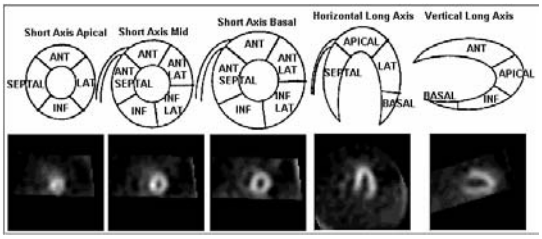
- SPECT (Single Photon Emission Computed Tomography)
- Two 3D sets of images, backprojected into six 2D sets of images
- Provides information about the left ventricle perfusion only



- Low resolution (64x64), b&w
- 2 series (rest & stress study)

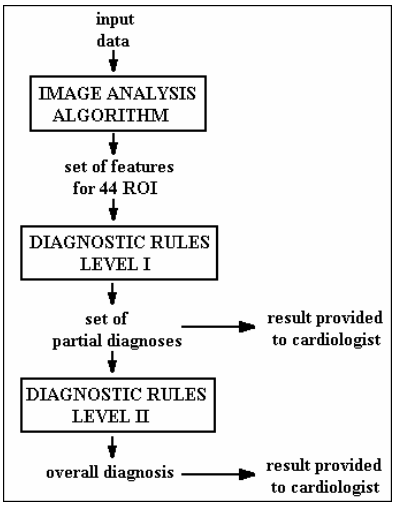


- 22 defined regions of the heart LV muscle (partial diagnoses)
- Five images for each study are chosen from 2D sets



SPECT Heart Image Analysis System

- GOAL: semi-automation of cardiac SPECT diagnostic process
- Image analysis in combination with ML used to mimic a diagnostic process performed by a cardiologist
- Image analysis identifies key features from cardiac SPECT images
  - 10 images (stress and rest studies)
  - patient sex
- Next, two types of diagnostic rules are generated
  - for partial diagnoses
    - input: features extracted from SPECT images
    - rules generated heuristically
  - for the overall diagnosis using
    - input: partial diagnoses
    - rules generated by the CLIP3 ML algorithm



SPECT Heart Image Analysis System

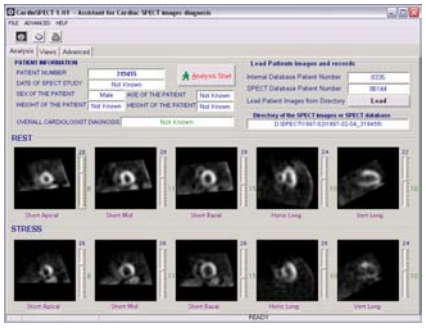
Results for partial diagnoses

	[%]	Correct / Total
Overall diagnosis: NL (49 pat.)	89.52	965 / 1078
Overall diagnosis: IS (27 pat.)	82.32	489 / 594
Overall diagnosis: INF (47 pat.)	81.53	843 / 1034
Overall diagnosis: ART (31 pat.)	81.52	556 / 682
Overall diagnosis: IS-IN (48 pat.)	77.65	820 / 1056
ENTIRE NEW DB (267 pat.)	81.34	4778 / 5874

Results for overall diagnoses

Sensitivity		Specificity		Predictive Accuracy		# rules
[%]	Correct / Total	[%]	Correct / Total	[%]	Correct / Total	
80.00	12 / 15	84.30	145 / 172	83.96	157 / 187	3

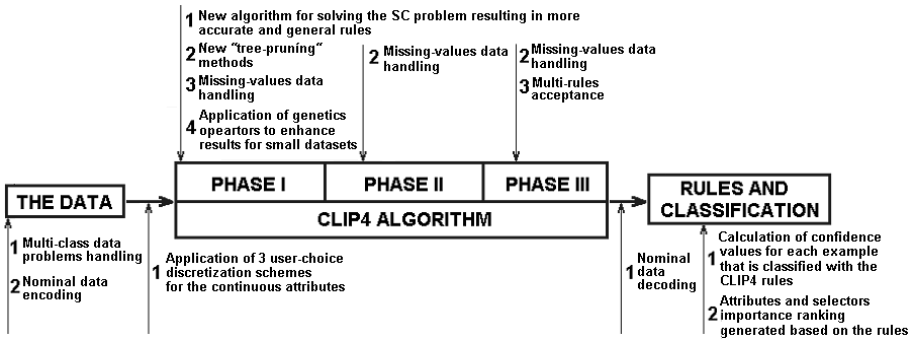
- Simple, easy to understand, and highly accurate diagnostic rules
- System can be used as an assistant tool by cardiologists to help them to make more consistent diagnosis of cardiac SPECT studies



UCD/HSC Center for Computational Biology Poster Session, March 8, 2002 Lukasz Kurgan, 2002

CLIP4 Algorithm

- The algorithm
  - partitions data into subsets using a tree structure and then generates production rules only from subsets stored at the leaf nodes
  - generates model of the data that consists of
    - well-generalized classification rules
    - ranked attributes and selectors



Ensemble of CLIP4 Classifiers

- An ensemble of classifiers is generated by injecting randomness into the CLIP4 algorithm
  - Multiple, different classifiers are generated
  - The classifiers are combined using weighted voting scheme

- The ensemble of CLIP4 classifiers algorithm

Given: Training data, where $x_j \in X$ is an example, and $y_i \in Y$ is its corresponding class label

Initialize: Set the semi-random mode of solving the SC problem, T - the number of the classifiers to be generated

For t = 1, … ,T

1. **Generate a set of rules (classifier) using the CLIP4 algorithm**
2. **Assign a weight α_t to the generated classifier $c_i : X \rightarrow Y$**

   where: α_t = accuracy of c_i on the entire training data set

Classify the examples using:

$$y_i = \arg\max_i \left(\sum_{j=1}^{T} \alpha_j c_j \; if \; c_j = y_i\right)$$

Ensemble of CLIP4 Classifiers

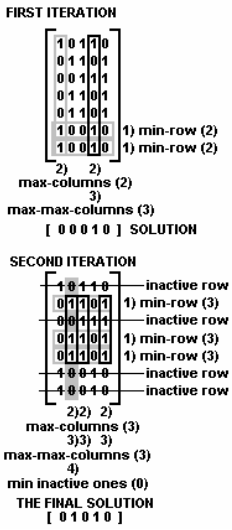
- **Ensemble classifier generation is based on the semi-random method for obtaining solutions for the set covering problem**

  Given: **BINary matrix**

  Initialize: **Remove all empty rows from the BINary matrix; if the matrix has no ones then return error**

  1. **Select active rows that have the minimum number of 1's in rows – min-rows**
  2. **Select the columns that have the maximum number of 1's within the *min-rows – max-columns***
  3. **Select randomly one of the *max-columns***
  4. **Add the selected column to the solution**
  5. **Mark the inactive rows, if all the rows are inactive then terminate, otherwise go to 1**

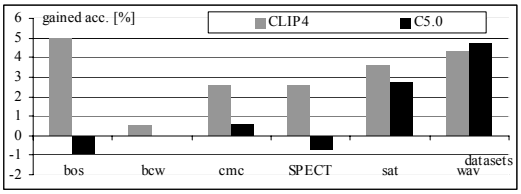
- **Each classifier is generated separately using the same training data set**



Ensemble of CLIP4 Classifiers

Dataset	CLIP4 with ensemble accuracy. [%]	CLIP4 with ensemble max. accuracy [%]	CLIP4 single classifier accuracy [%]	Gained accuracy [%]
SPECT	88.7	90.4	86.1	2.6
sat	83.1	84.0	79.5	3.6
wav	78.8	78.8	74.5	4.3
bos	75.5	N/A	70.5	5.0
bcw	95.7	N/A	95.2	0.5
cmc	49.5	N/A	46.9	2.6

- **Significant improvement, when comparing to the accuracy of a single CLIP4 classifier**
  - **average achieved gain of accuracy was about 3.1%, which corresponds to about 15% decrease of the error rate**
- **Testing results show that a significant gain of accuracy is achieved for the ensemble of 3 to 4 classifiers**
  - **high usefulness of the proposed method**



- **Increase of accuracy is achieved for all datasets**
  - **in contrary to boosting technique**

UCD/HSC Center for Computational Biology Poster Session, March 8, 2002 Lukasz Kurgan, 2002

Ensemble of CLIP4 Classifiers and SPECT Heart Image Analysis System

- **Using the ensemble of classifiers method, the accuracy of the rules used to generate overall diagnosis for the diagnostic system was improved**
  - **a single CLIP4 classifier achieved accuracy of 86.1%**
  - **the ensemble of 4 CLIP4 classifier achieved accuracy of 90.4%**
  - **the gained 4% accuracy reduces the previous error rate by 31%**
    - **significant and essential improvement for the entire medical diagnostic system**

References

Cios, K. J., Wedding, D. K., and Liu, N., CLIP3: Cover Learning Using Integer Programming. *Kybernetes*, 26 (4-5), pp. 513-536 (The Norbert Wiener 1997 Outstanding Paper Award), 1997

Cios K. J. & Kurgan L., Hybrid Inductive Machine Learning: An Overview of CLIP Algorithms, In: Jain L.C., and Kacprzyk J. (Eds.) *New Learning Paradigms in Soft Computing,* pp. 276-322, Physica-Verlag (Springer), 2001

Kurgan, L.A., Cios, K.J., Tadeusiewicz, R., Ogiela, M. & Goodenday, L.S., Knowledge Discovery Approach to Automated Cardiac SPECT Diagnosis, *Artificial Intelligence in Medicine*, vol. 23/2, pp.149-169, 2001

Kurgan L. & Cios K.J., Ensemble of Classifiers to Improve Accuracy of the CLIP4 Machine Learning Algorithm, accepted for the SPIE's International Symposium on Sensor Fusion: Architectures, Algorithms, and Applications VI, 2002