ANNs and coccolithophores
The *COGNIS* system

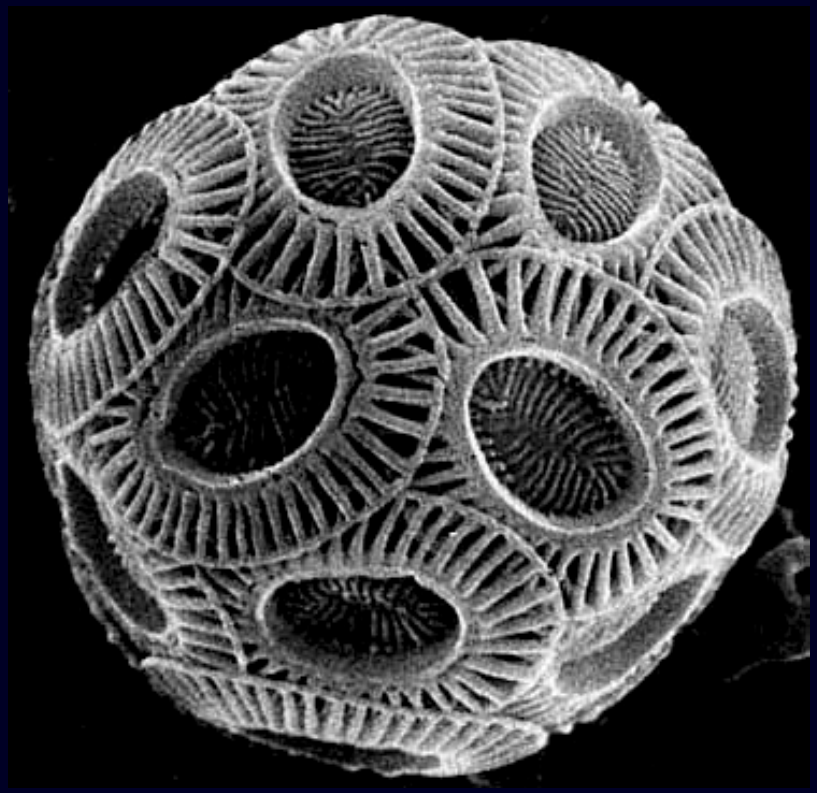
Patrick Quinn
Jörg Bollmann
Miguel Vela
Ralph Schiebel
Hans Thierstein

ETHZ Zürich

What are coccolithophores?

- Single-celled, calcifying marine algae
- Widespread distribution in surface waters, very abundant (*E. huxleyi* blooms)
- Important primary producers, crucial to global carbon cycle, produce DMS
- High biodiversity and useful to study genetic processes

What are coccolithophores?

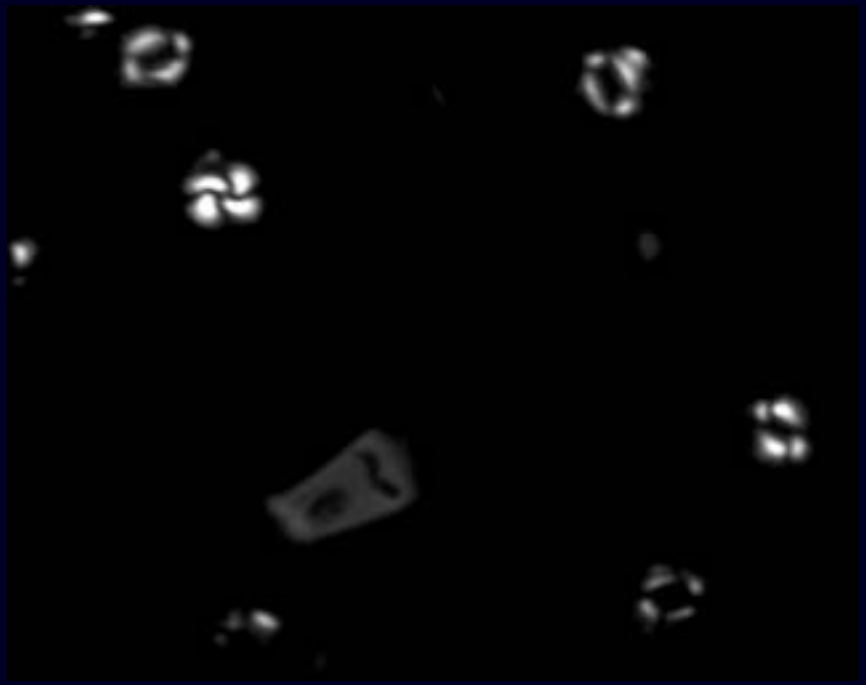


E. huxleyi coccosphere - size 5µm

Fossil coccolithophores

- Coccolithophores have 200 Ma geolgical record
- Usually found as isolated coccoliths and are classified within *calcareous nannofossils*
- Useful for dating rocks (oil industry), geologic and oceanographic interpretation
- Taxa identified by morphology of coccoliths in LM and SEM

Fossil coccolithophores



Fossil coccoliths in LM - size 2 µm

Motivation

- Analysis of nannofossils is tedious and time consuming
- Goal is to speed up process of counting and be able to search for rare coccoliths without excess manpower
- Automation will produce more reliable quantitative data leading to better interpretations

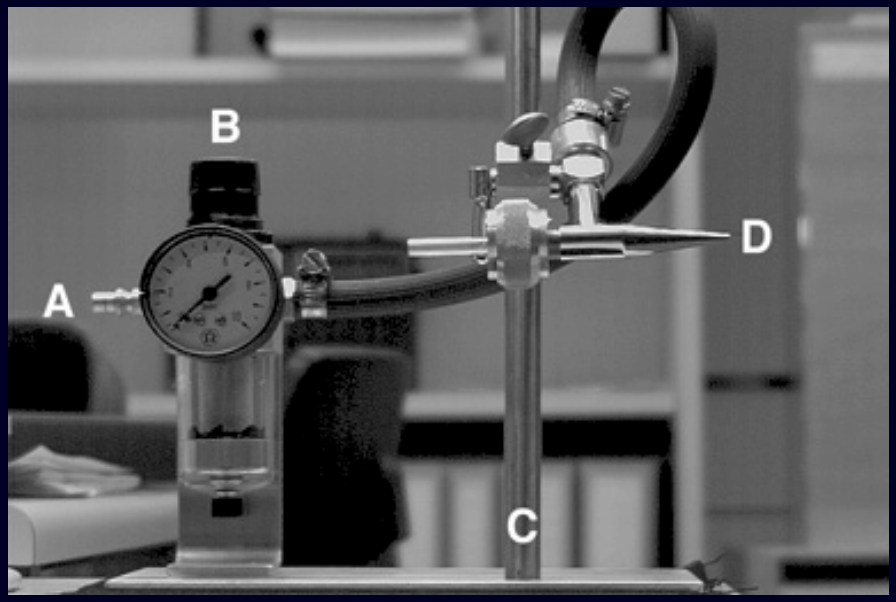
Areas of development

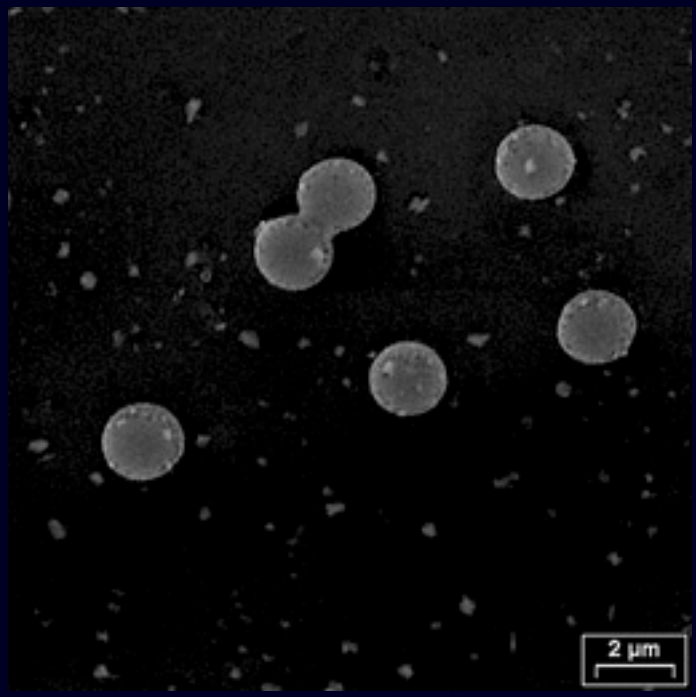
- To achieve our aim of producing a computer guided nannofossil identification system (*COGNIS*) we have made developments in 3 areas:
    1. Sample preparation
    2. Automated microscopy
    3. Artificial Neural Networks

Sample preparation

- The success of automated collection and NN classification is dependent upon good sample preparation - *what you get out depends on what you put in*
- Evenly distributed, thin slides with well dispersed *microbeads* for calibration can now be produced using our *SMS* technique

Sample preparation





SMS technique - spray gun and microbeads

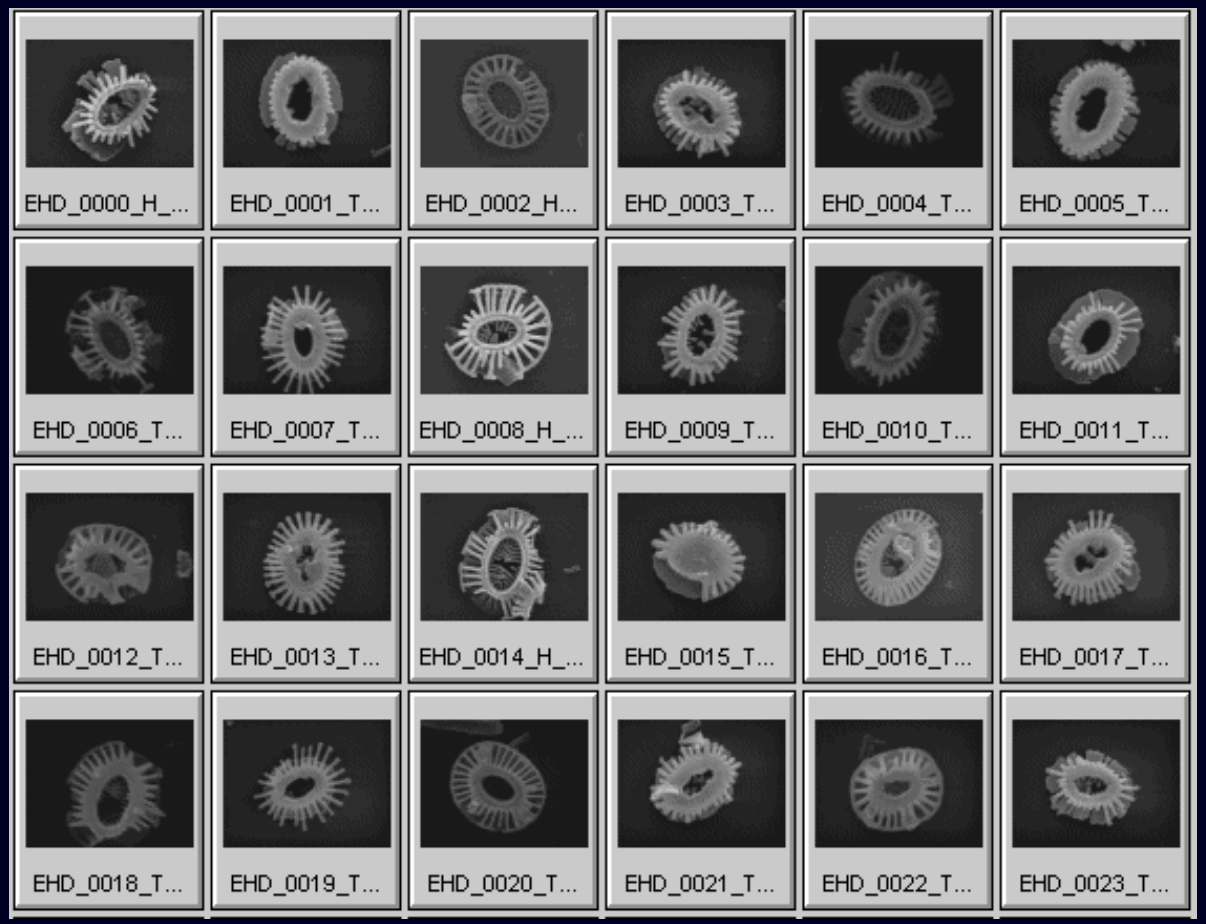
Automated microscopy

- Aim is rapid collection and storage of single microfossil particles
- Requires mechanical stage, autofocus, object detection and segmentation
- In collaboration with *Leica* and *Philips* we have developed 3 systems - SEM, Transmitted LM and Incident LM (Forams)

Automated microscopy

Automated Philips XL30 SEM

Automated microscopy

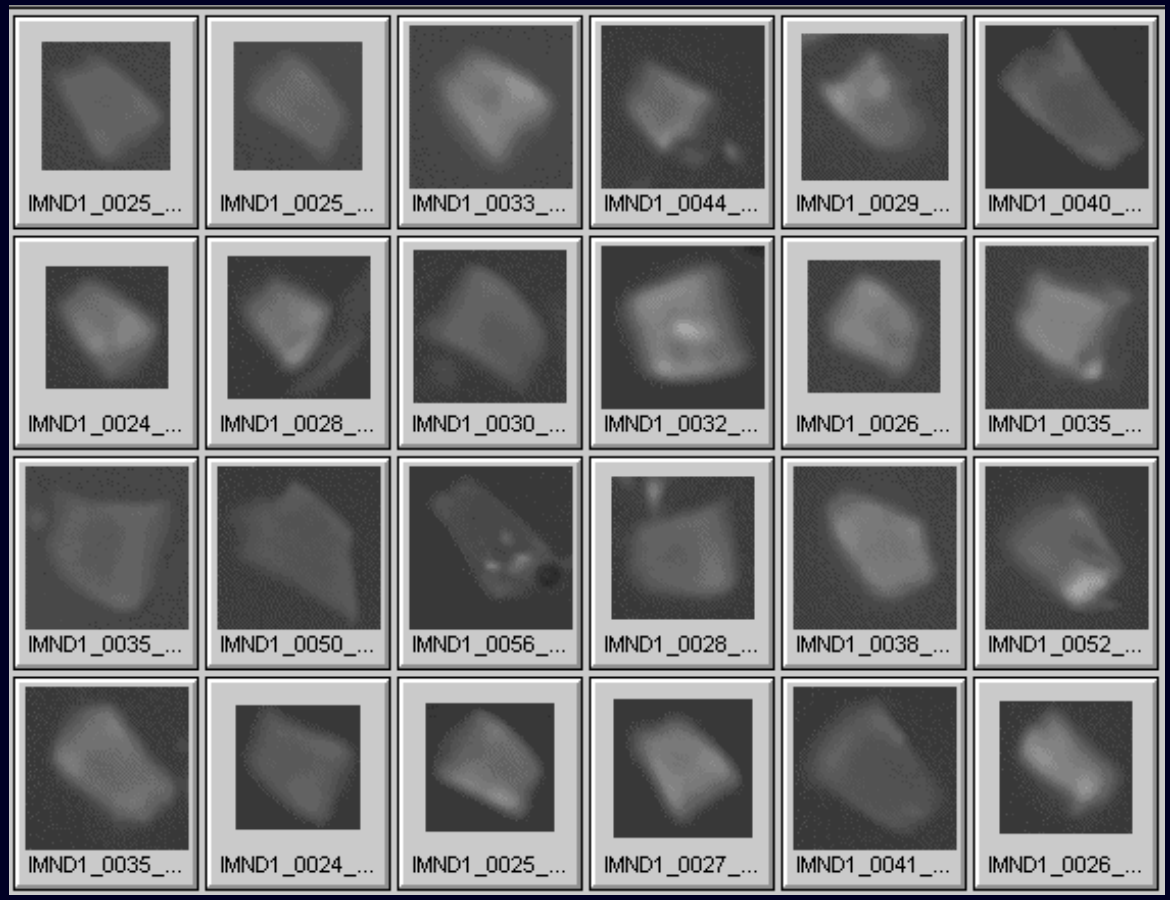


Images collected automatically in the SEM

Automated microscopy

Leica DMRXA transmitted LM

Automated microscopy



Images collected automatically in the LM

ANNs

- Nannofossils are well-suited to image analysis:
  - Usually flat 2D images
  - Can be converted to greyscale without losing information
  - Species recognition is based on purely morphologic criteria including outline of coccolith and nature of several concentric regions

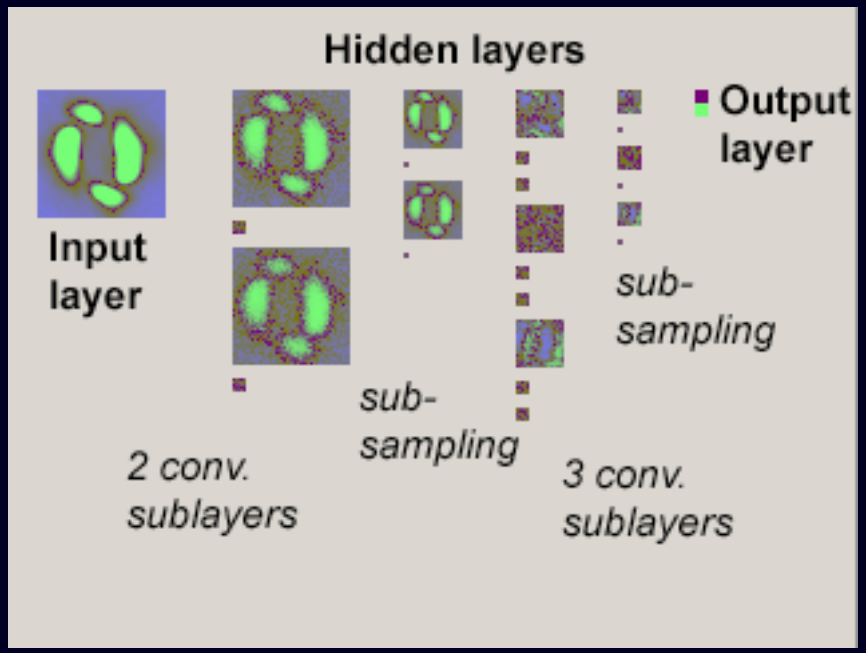
ANNs

- Nannofossils are particularly suited to traditional image analysis - *shape analysis pattern recognition and feature extraction*
- But we need to identify many different types of objects in each sample
- We therefore opted for a Neural Network approach

ANNs

- In order to deal with some shift and distortion in the nannofossil images we use c*onvolutional* rather than fully-connected NNs
- These use:
    1. Local receptive fields
    2. Shared weights
    3. Subsampling

ANNs

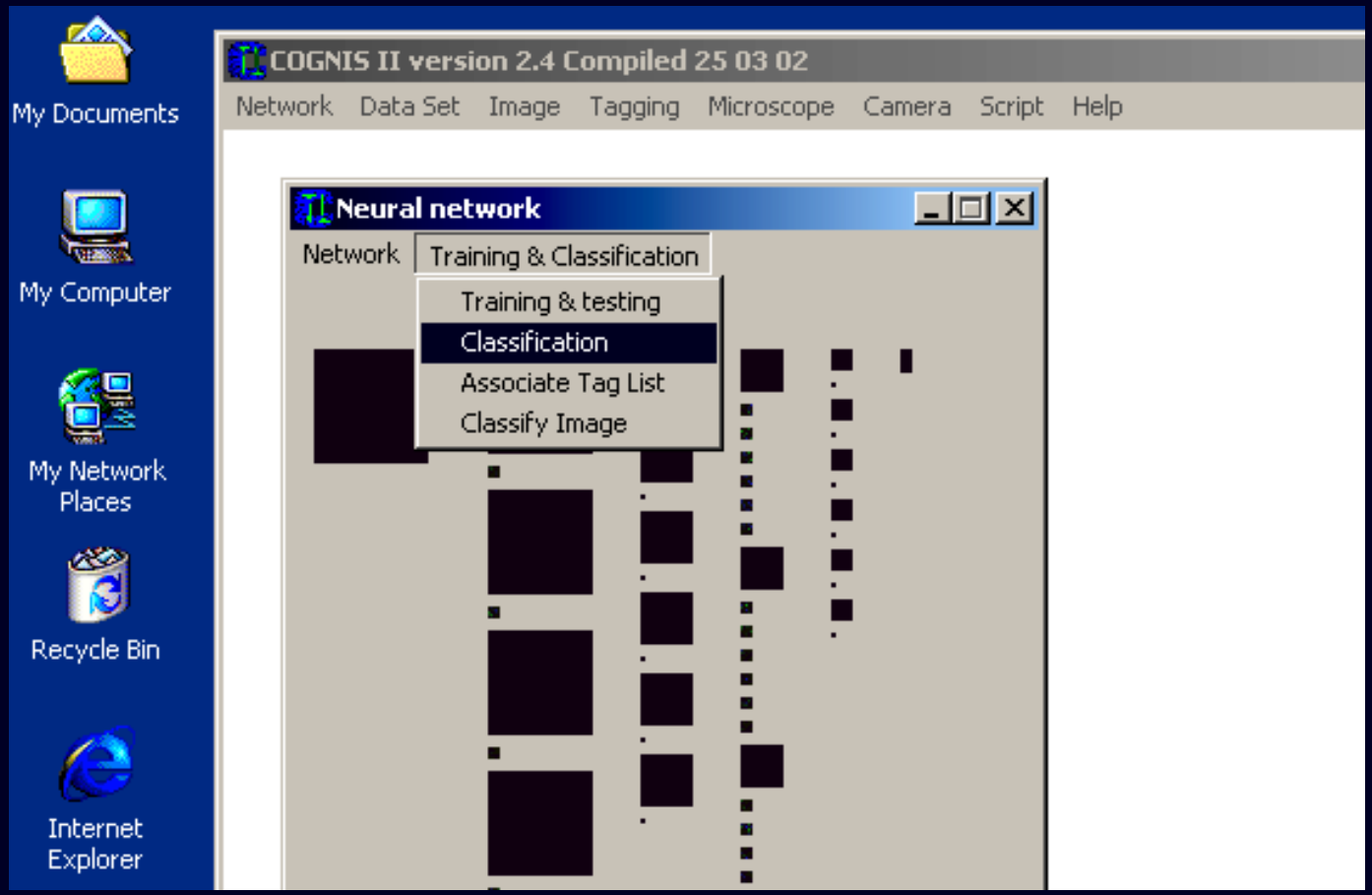


A Convolutional NN

ANNs

- Developed Windows-Based *COGNIS* program with versatile user-friendly Neural Networks
- But *COGNIS* is a lot more:
  - Controls microscope
  - Processes, archives and tags images

ANNs



COGNIS Windows-based program

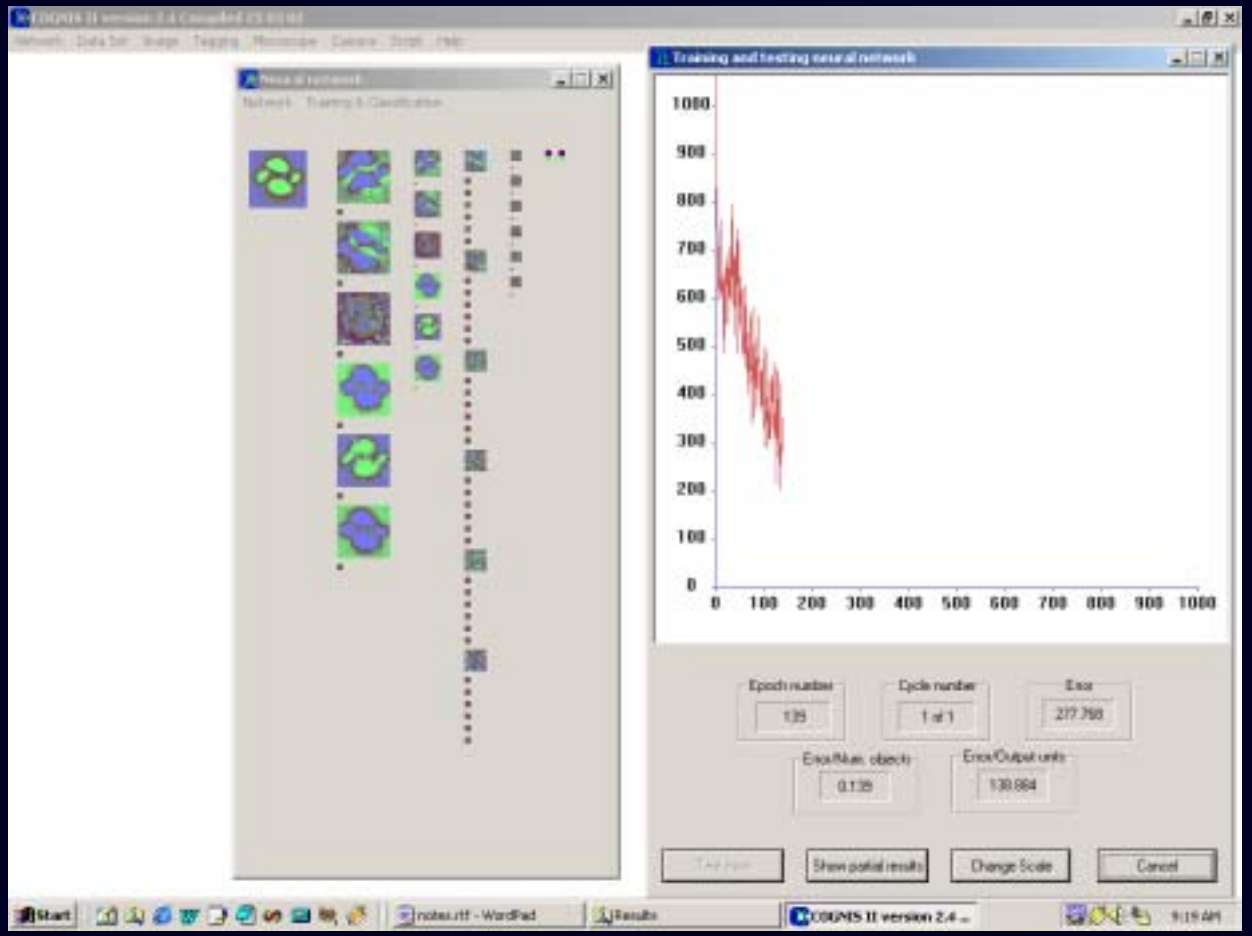
ANNs

- The *COGNIS* program gives the user full control to build fully connected or convolutional NNs of any architecture
- These can deal with greyscale images or numerical data
- All training parameters - *weights*, *bias*, *learning rate*, *epochs* - can be adjusted to produce the optimal NN for each application

ANNs

- Training is carried out using a data set file (.dsf) constructed from folders containing known images
- The NN learns by *backpropagation* - the output class is compared to the true class, an error is calculated and the weights in the network are adjusted accordingly

ANNs



Training a neural network with *COGNIS*

ANNs

- Knowing what NN structure to use for what application requires experience and practise
- Tested the program with the *Olivetti Research Laboratory Face Database* - 10 poses of 40 people
- Best net produced 85% correct classification

ETH

cognisII

ANNs

The ORL Face Database

ANNs

- Nannofossil images are far more varied than the ORL faces and a dataset with >100 images per class is usually needed
- Despite the tolerance of convolutional NNs to variation, we preprocess our nannofossil images:
    - Rotate objects to homologous position
    - Convert to greyscale

ANNs

- Classification rate is highly dependent upon the number and type of classes which vary depending upon the task
- Potential tasks range from:

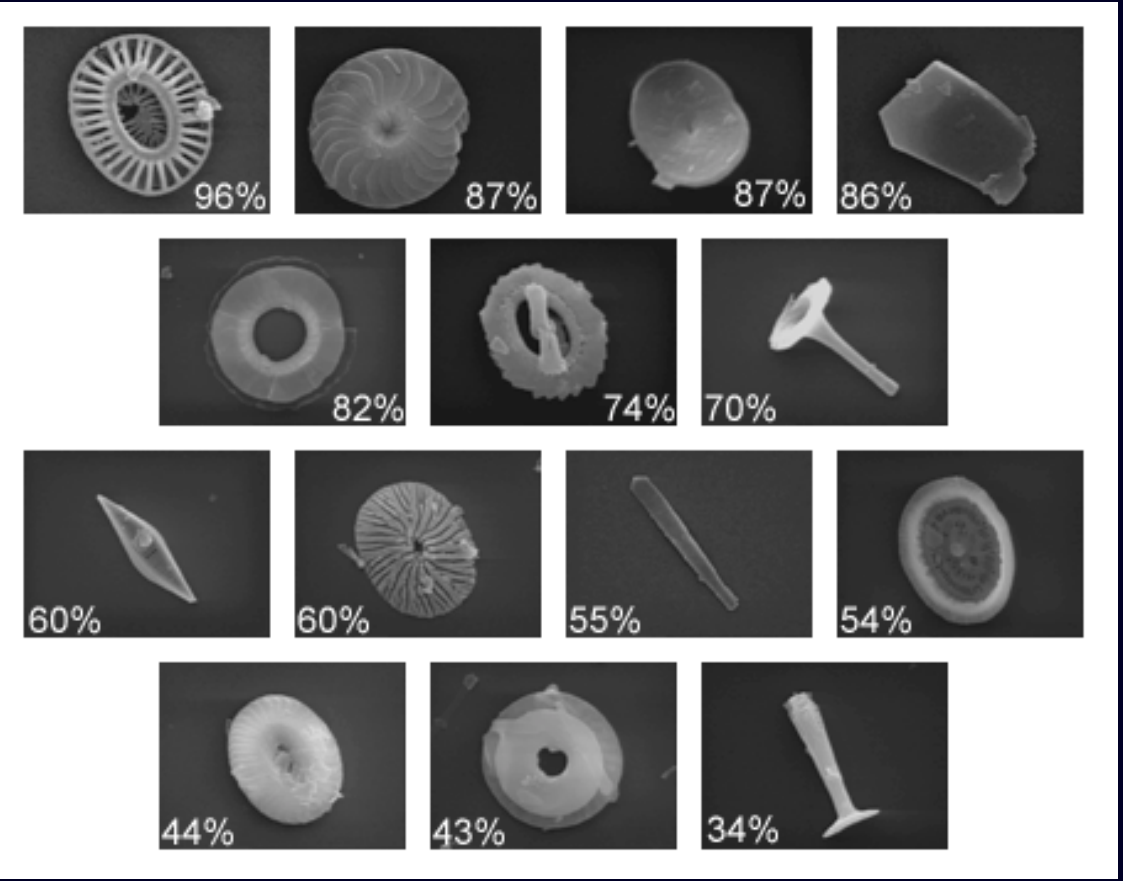
  - Classification of all particles in a sample, to

  - Identification of a single type of particle among all others in a sample

ANNs

- *Example 1*. SEM classification of a well-preserved Holocene sample:
  - Using 5 layered NN with 48x48 input layer
  - Training with 14 classes - 100 image per class
  - Average classification rate of 82%
  - Individual classification rates varied from 96% to 34%

ANNs

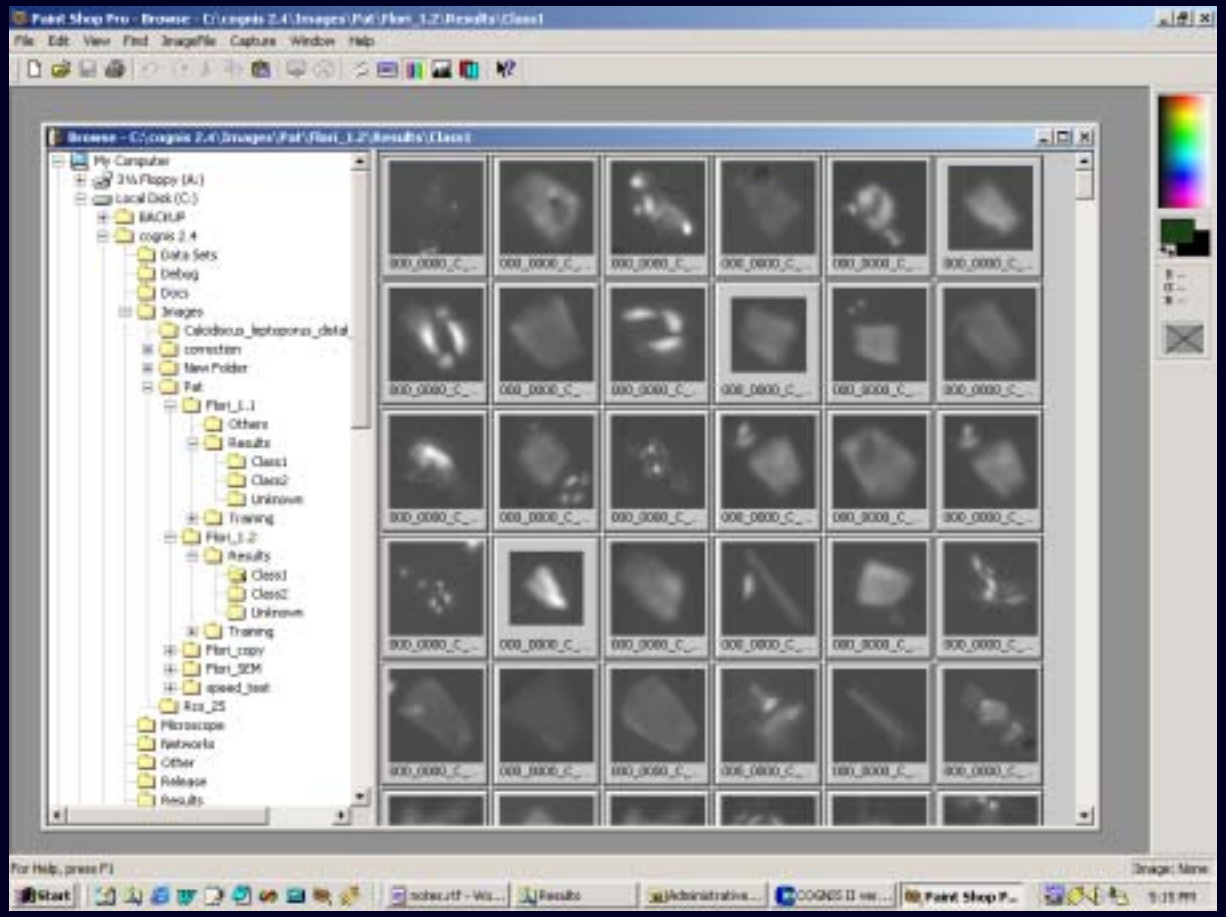


Classification results

ANNs

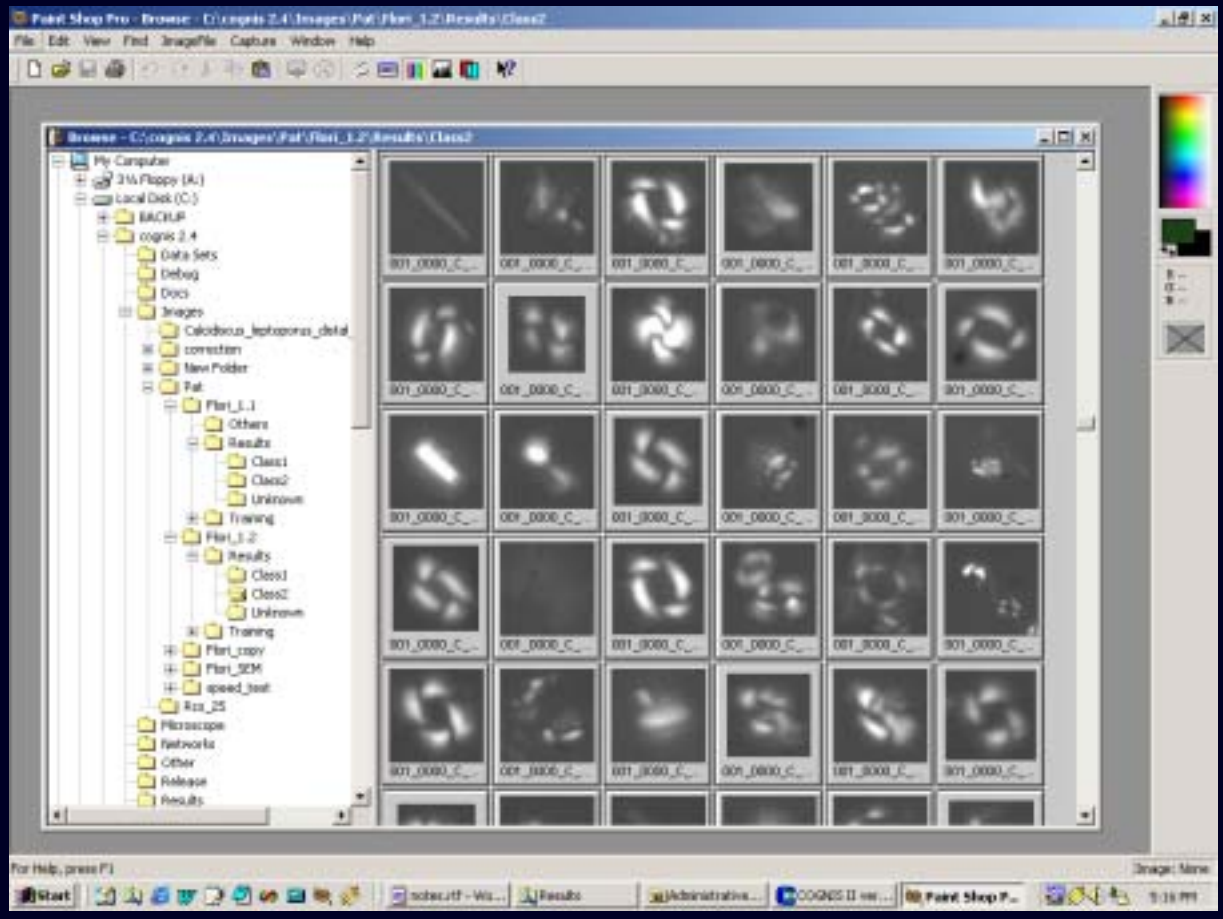
- *Example 2*. LM detection of single particle in well-preserved Holocene sample
  - Same 5 layered NN with 48x48 input layer
  - 2 classes with 1000 images
  - 98% of the desired particles correctly identified
  - Some other things put in with them

ANNs



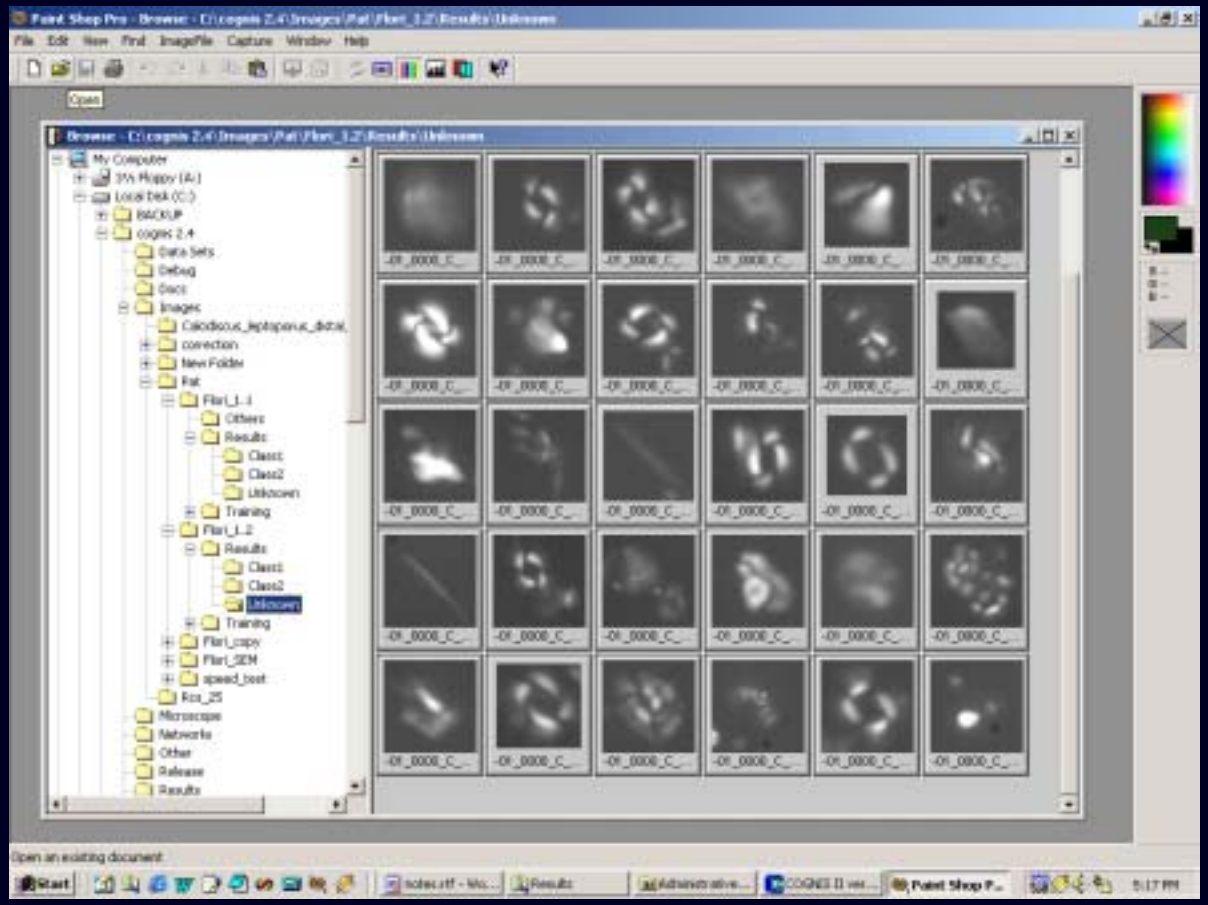
Class 1

ANNs



Class 2

ANNs

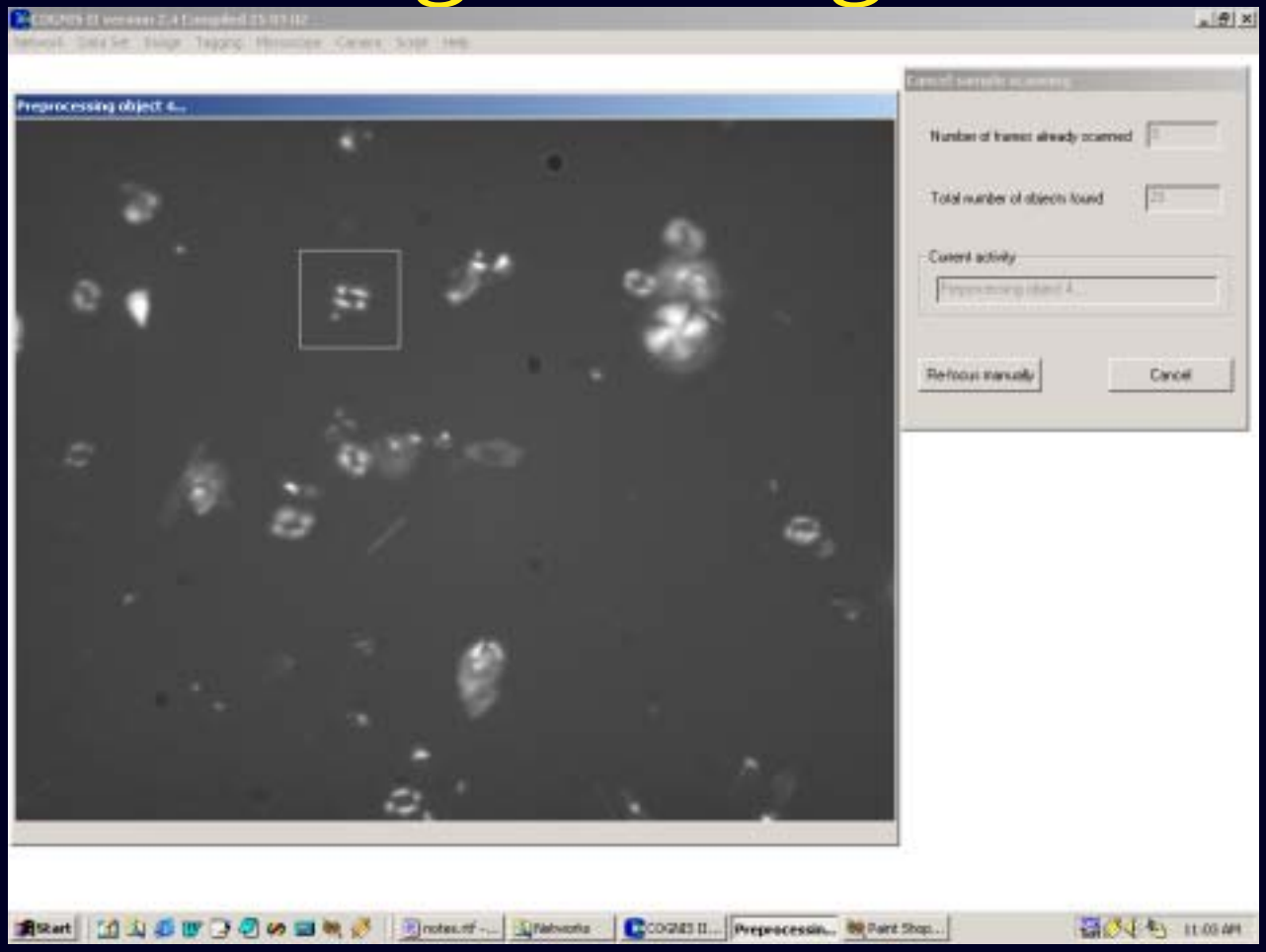


Unknown

Putting it all together

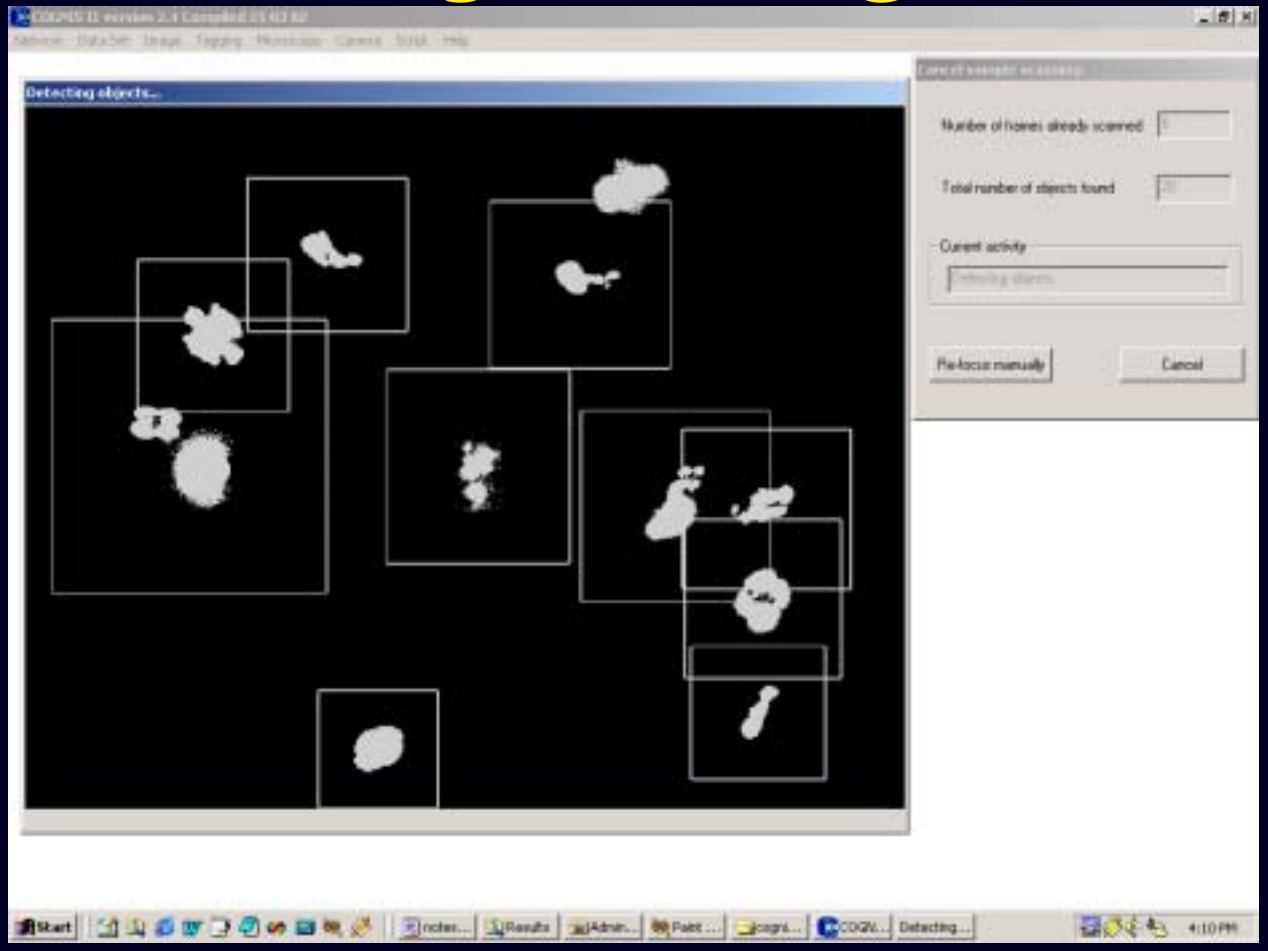
- Latest version of *COGNIS* permits on-line classification
- Simultaneous collection, preprocessing and classification of images
- Current maximum working speed of c. 500 objects per hour

Putting it all together



Online collection and classification

Putting it all together



Online collection and classification

The future

- Aim to improve preparation, efficiency of machine and refine NN success
- Will be used in first major research project this year
- Carrying out tests in collaboration with major oil company