ieeta instituto de engenharia electrónica e telemática de aveiro
Universidade de Aveiro

universidade de aveiro

IBERGRID
4rd IBERIAN GRID INFRASTRUCTURE CONFERENCE
Braga (Portugal) May 24-27, 2010

Automated endoscopic capsule analysis using a Grid computing environment

Ilídio C Oliveira, Luís Alves, Eduardo Dias, David Pacheco, Sérgio Lima
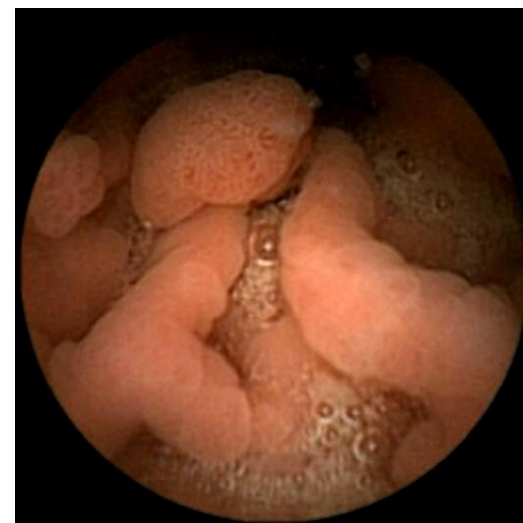*U. Aveiro & IEETA*
João Barros, Miguel P Monteiro, Jorge A Silva
INEB & FEUP
José Maria Fernandes, J P S Cunha, António Sousa Pereira
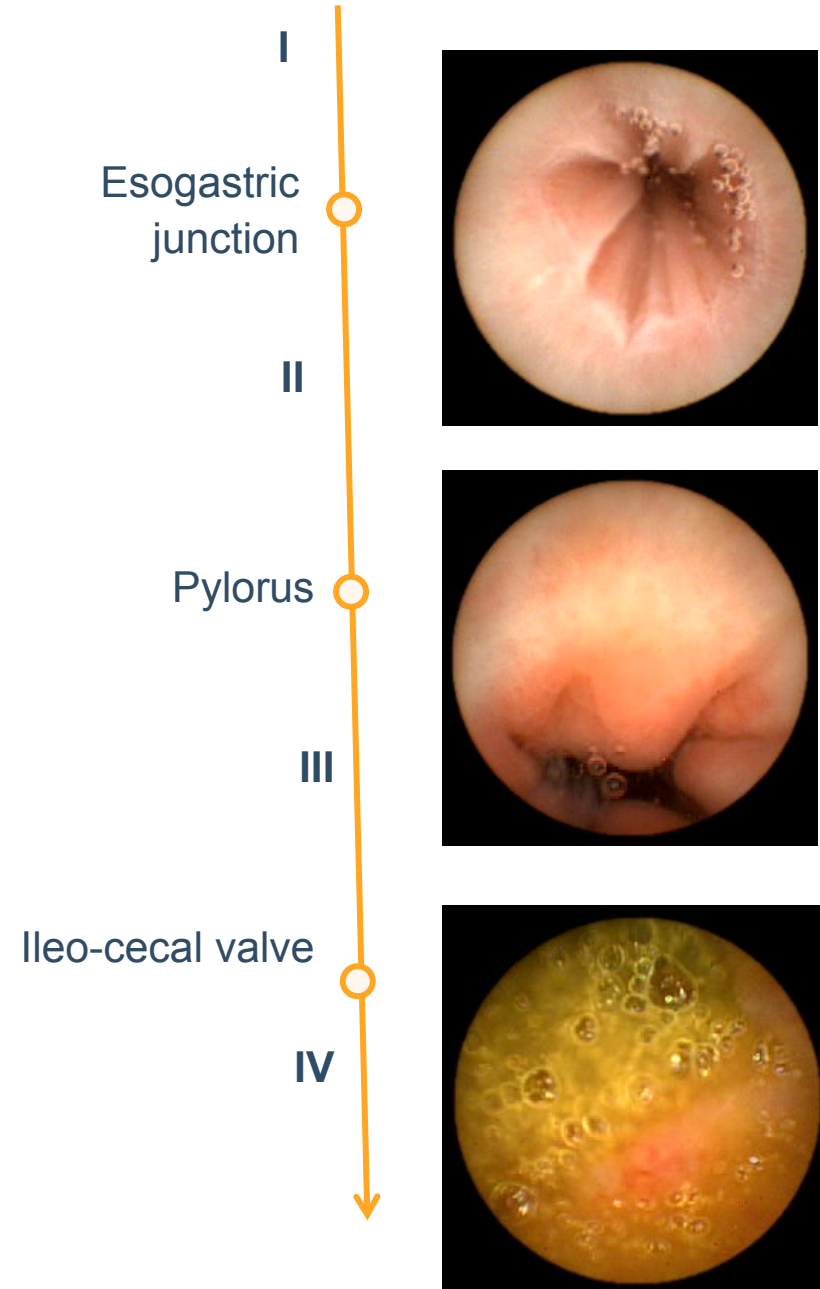*U.Aveiro & IEETA*

http://www.ieeta.pt/sias

Wireless Endoscopic Capsule (WEC) is emerging in diagnosis

- 6 to 8 hours long video

- 2 hours to review

- automated image processing methods can help
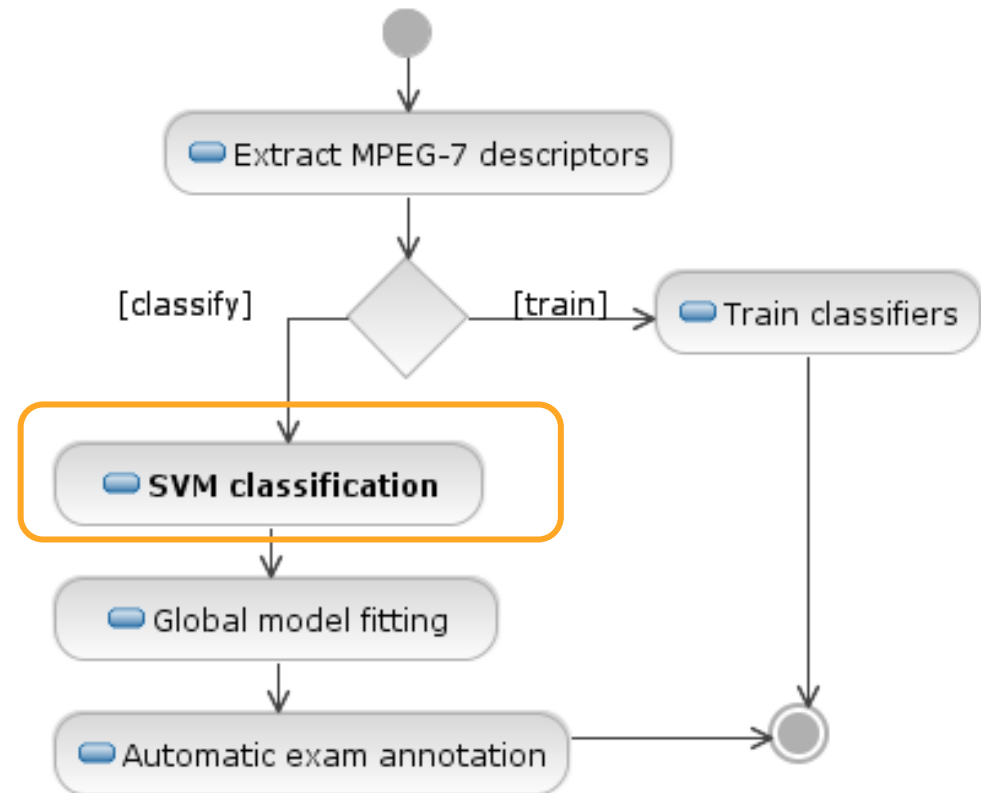 - CADe: polyps, bleeding,...

ieeta

Automated Topographic Segmentation (ATS) can reduce review times

- Locate the four main topographic regions

- Uses Support Vector Machine (SVM) classification applied to MPEG-7 descriptor vectors

  ▸ http://dx.doi.org/10.1109/TMI.2007.901430

- ATS deployed in CapView visualization tool

  ▸ www.capview.org

- Usually takes >1h

I

Esogastric junction

II

Pylorus

III

Ileo-cecal valve

IV

The Grid as a "cloud" for capsule R&D and clinical practice?
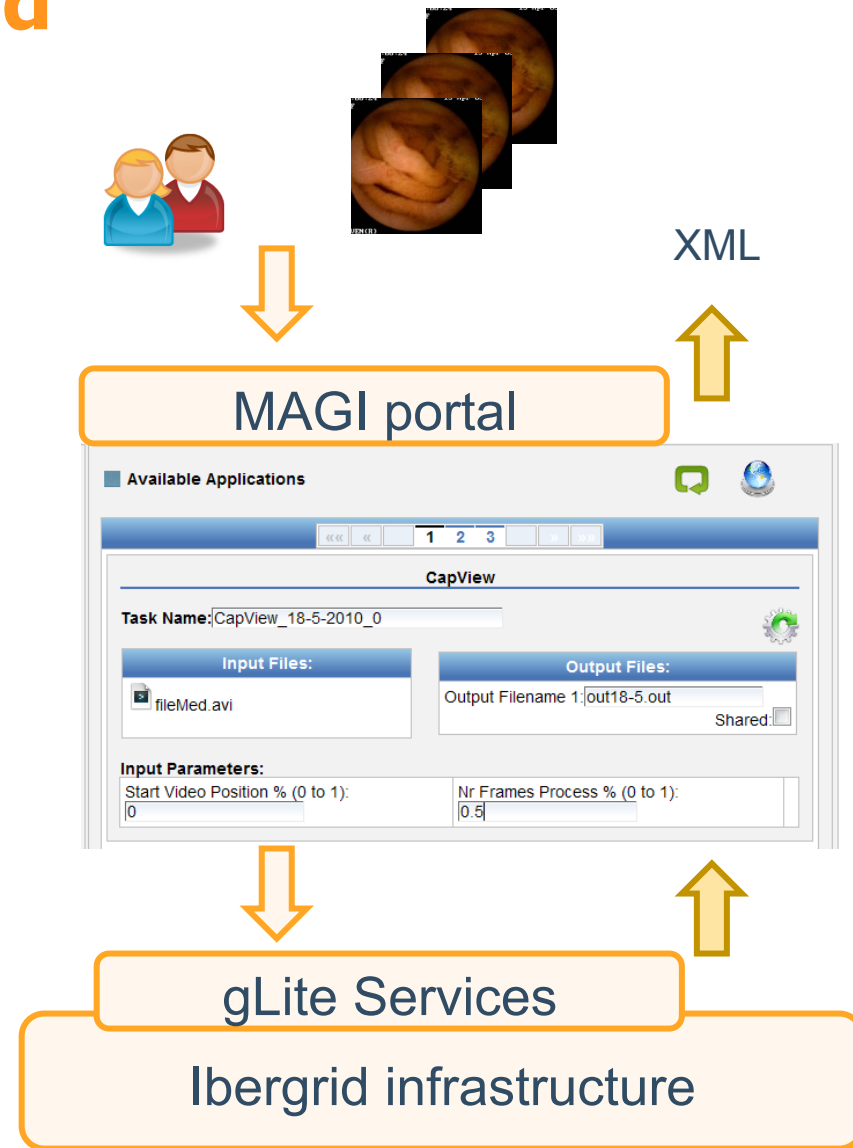
- Repositories of scientific data
  - ▶ GeresMed Project
  - ▶ Data sets anonymized
- Processing videos
  - ▶ Run SVM Classification in Grid infrastructures
  - ▶ Run existing algorithm "as is"
  - ▶ Domain partition is feasible

Extract MPEG-7 descriptors

[classify] [train] → Train classifiers

SVM classification

Global model fitting
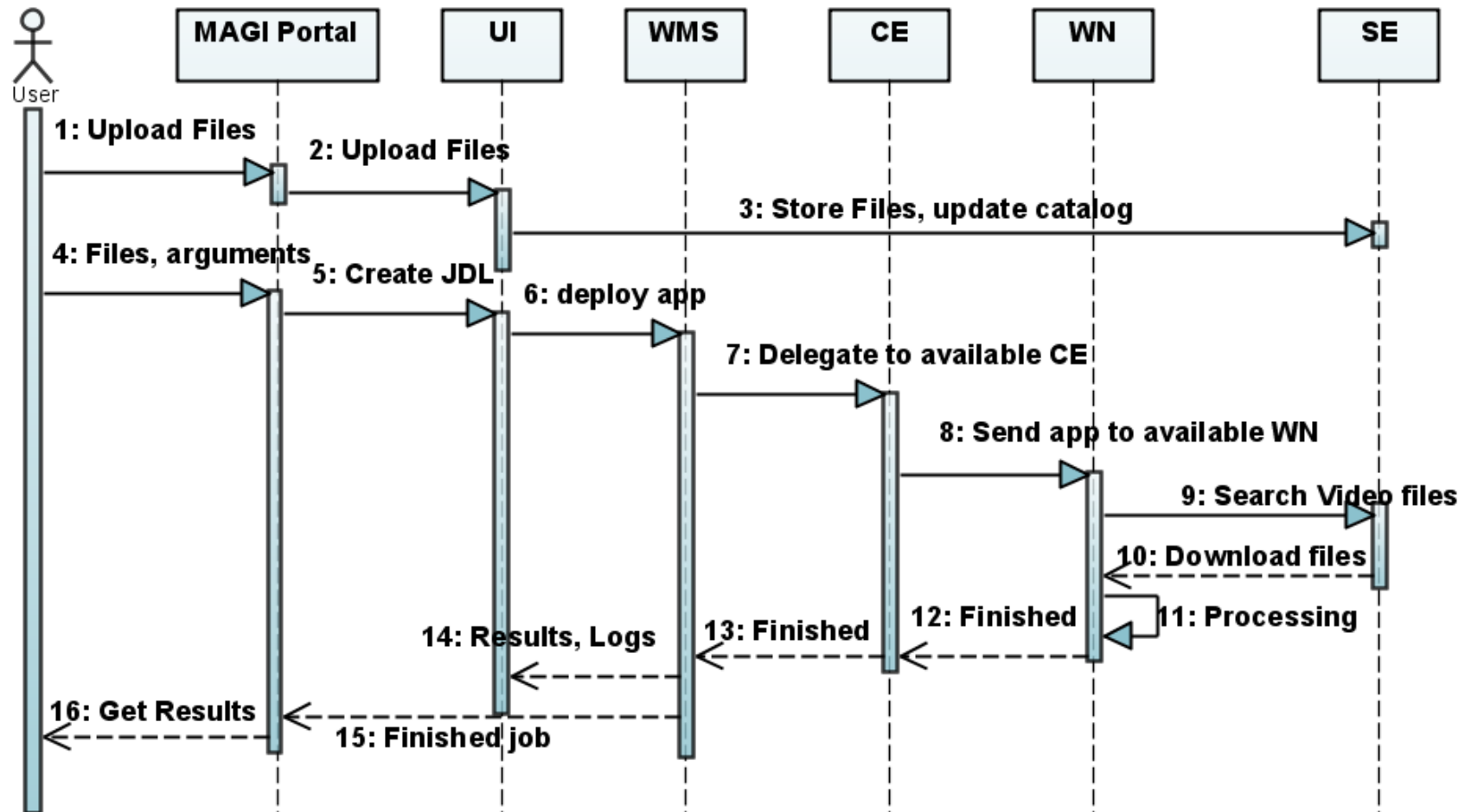
Automatic exam annotation

ieeta

Enabling friendly Grid analysis

1. Access the MAGI Grid interfacing portal
2. Upload data into MAGI & select the "CapView" app.
3. Data and jobs description submitted to the Grid
4. The progress of jobs is monitored by MAGI and output results aggregated
5. Classifications calculated in the Grid used in the segmentation step

XML

MAGI portal

Available Applications

| «« | « | **1** | 2 | 3 | | » | »» |

CapView

Task Name: CapView_18-5-2010_0

Input Files:	Output Files:
fileMed.avi	Output Filename 1: out18-5.out
	Shared: ☐

Input Parameters:

Start Video Position % (0 to 1):	Nr Frames Process % (0 to 1):
0	0.5

gLite Services

Ibergrid infrastructure

ieeta

Communication between modules

Available Applications

| | 1 | 2 | 3 | | |

CapView

Task Name: CapView_18-5-2010_0

Input Files:
fileMed.avi

Output Files:
Output Filename 1: out18-5.out
Shared:

Input Parameters:

| Start Video Position % (0 to 1): | Nr Frames Process % (0 to 1): |
| 0 | |

View/Search Data

| My Private Data | Public Data Pool | Task Results | Organize Data by Studies... | My Processing Tasks |

Processing Tasks

Task Name ⇕	Starting Date ⇕	Current Status ⇕	Finish Status ⇕	Running Time ⇕	Action
CapView_16-4-2010_0	2010-04-16 13:57:13.418	SCHEDULED			✖
CapView_16-4-2010_0	2010-04-16 13:56:51.24	RUNNING			✖
CapView_16-4-2010_0	2010-04-16 13:56:23.747	SCHEDULED			✖
CapView_16-4-2010_0	2010-04-16 13:54:56.803	RUNNING			✖
CapView_19-3-2010_0	2010-03-19 01:15:58.041	FINISHED WITH SUCCESS	FINISHED WITH SUCCESS	00:57:20	⬇
CapView_19-3-2010_0	2010-03-19 01:15:17.858	FINISHED WITH SUCCESS	FINISHED WITH SUCCESS	00:58:01	⬇
CapView_19-3-2010_0	2010-03-19 01:14:35.272	FINISHED WITH SUCCESS	FINISHED WITH SUCCESS	01:04:39	⬇
CapView_19-3-2010_0	2010-03-19 01:13:47.332	FINISHED WITH SUCCESS	FINISHED WITH SUCCESS	00:49:28	⬇

Desktop vs Grid, no paralellization

Exam size		Ex1	Ex2	Ex3
KB		210.158	408.686	659.098
Desktop (with GUI)	Elapsed	00:32:37	01:03:38	01:30:52
Desktop (no GUI)	Elapsed	00:16:10	00:31:00	00:44:22
Grid processing time (GPT)	Start	13:08:47	10:29:17	13:08:50
	End	13:30:15	11:10:38	14:10:07
	Elapsed	**00:21:28**	**00:41:21**	**01:01:17**
MAGI Round-Trip (MAGI-RT)	Start	13:05:12	10:23:50	13:04:26
	End	13:49:14	11:48:55	14:38:20
	Elapsed	**00:44:02**	**01:25:05**	**01:33:54**
Grid "overhead" (MAGI-RT - GPT)		00:22:34	00:43:44	00:32:37
% of MAGI-RT		51%	51%	35%

ieeta

No multiplicative factor in the Grid as the data is partioned

Strategy	Run #1	Run #2	Run#3	Run #4	Cluster
1-fold	00:41:21	00:40:13	00:38:02	00:39:05	01:01:06
2-fold	01:14:38	01:03:44	02:18:27	00:25:14	00:30:53
4-fold	01:07:31	00:25:17	02:04:18	00:27:14	00:15:43
8-fold	01:19:29	01:10:53	00:52:30	01:37:44	00:08:13

Ibergrid   EGEE   IBM BladeCenter JS21, with

PowerPC 970@ 2.3 GHz

ieeta

High variability in computing times

	4-fold partition				8-fold partition							
Run #1	@VO BING, ce04.ncg.ingrid.pt											
Start	13:53:00	13:53:00	13:53:01	13:53:01	15:58:08	15:58:04	15:58:07	15:58:10	16:13:1	16:03:04	16:03:05	16:03:06
End	14:26:31	15:00:3	14:56:	14:26:21	16:03:20	16:29:44	16:13:23	16:19:10	17:17:	16:19:15	16:13:29	16:07:45
Elapsed	**00:33:31**	**01:07:31**	**01:03:51**	**00:33:20**	**00:05:12**	**00:31:40**	**00:15:16**	**00:21:00**	**01:04:32**	**00:16:11**	**00:10:24**	**00:04:39**
Avg	00:49:33		*Min*	00:33:20	*Avg*	00:21:07					*Min*	00:04:39
Std	00:18:41		*Max*	**01:07:31**	*Std*	00:19:37					*Max*	**01:04:32**
Run #2	@VO BING, ce04.ncg.ingrid.pt											
Start	11:27:49	11:27:48	11:27:49	11:42:43	15:54:05	15:58:33	15:58:26	15:58:2	15:58:19	15:58:32	15:58:29	15:58:31
End	11:38:44	11:51:27	11:47:19	11:53:05	15:59:48	16:09:52	16:47:24	17:04:	17:03:32	16:14:03	16:31:46	16:15:27
Elapsed	**00:10:55**	**00:23:39**	**00:19:30**	**00:10:22**	**00:05:43**	**00:11:19**	**00:48:58**	**01:06:27**	**01:05:13**	**00:15:31**	**00:33:17**	**00:16:56**
Avg	00:16:06		*Min*	00:10:22	*Avg*	00:32:55					*Min*	00:05:43
Std	00:06:32		*Max*	**00:23:39**	*Std*	00:24:28					*Max*	**01:06:27**
Run #3	@VO BING, axon-g01.ieeta.pt and ce04.ncg.ingrid.pt											
Start	16:38:39	17:05:00	16:38:32	15:48:34	11:39:55	12:04:23	11:38:21	11:38:16	11:38:22	11:38:21	11:39:55	12:18:31
End	16:48:51	17:52:52	17:28:03	15:58:30	11:51:52	12:14:20	11:54:34	11:57:03	11:57:58	11:53:23	12:03:13	12:30:46
Elapsed	**00:10:12**	**00:47:52**	**00:49:31**	**00:09:56**	**00:11:57**	**00:09:57**	**00:16:13**	**00:18:47**	**00:19:36**	**00:15:02**	**00:23:18**	**00:12:15**
Avg	00:29:23		*Min*	00:09:56	*Avg*	00:15:53					*Min*	00:09:57
Std	00:22:19		*Max*	**00:49:31**	*Std*	00:04:30					*Max*	**00:23:18**
Run #4	@dteam EGEE, ce07.pic.es											
Start	02:01:07	02:08:54	02:02:22	02:02:23	02:15:24	02:15:27	02:17:08	01:19:56	02:17:20	01:19:56	02:47:26	02:47:23
Cluster	@IBM BladeCenter JS21 Dual Core 2.3 GHz 64-bit PowerPC 970											
Elsapsed	**00:15:43**	**00:15:42**	**00:15:39**	**00:15:35**	**00:08:13**	**00:08:11**	**00:08:13**	**00:08:10**	**00:08:11**	**00:08:10**	**00:08:12**	**00:07:59**
Avg	00:15:40		*Min*	00:15:35	*Avg*	00:08:10					*Min*	00:07:59
Std	00:00:04		*Max*	**00:15:43**	*Std*	00:00:05					*Max*	**00:08:13**

ieeta

Discussion

- Ibergrid/Ingrid is ready
- Grid nature includes variability
  - ▶ High variability in processing times
  - ▶ Significant "overhead" in job preparation
- Science requirements are best met
  - ▶ Assyncronous ("fire and forget")
  - ▶ Massive analysis
- Very good speedups with cluster paralellization
  - ▶ Potential for GPU approaches

Data Import

+ Add...	✕ Clear All

fileMed.avi
Done   Clear

Uploaded Files Info

fileMed.avi

File Name:	fileMed.avi	**Description:**	WEC Exam #1
Subject:	ProcID:N/D, Generic	**Modality:**	CapView capsule
Equipment:	Generic		

Clear Uploaded Data | Refresh

Private Folder	**Public Folder**

View/Search Data

My Private Data	Public Data Pool	Task Results	Organize Data by Studies...

Public Files

[   ] Search

Advanced Search

Modality:	Subject: ⇵	Date: ⇵	File Name: ⇵	
fMRI	John Doe	2010-01-06 1	job0.jdl	
Undefined	Generic	2010-01-07 1	videotest1.avi	
MRI	John Doe	2010-01-11 1	Screenshot-1.png	
Undefined	John Doe	2010-03-18 1	fileMed.avi	

Community View

0-05-23   13

Automated endoscopic capsule analysis using a Grid computing environment

Ilídio C Oliveira, Luís Alves, Eduardo Dias, David Pacheco, Sérgio Lima
U. Aveiro & IEETA
João Barros, Miguel P Monteiro, Jorge A Silva
INEB & FEUP
José Maria Fernandes, J P S Cunha, António Sousa Pereira
U.Aveiro & IEETA

http://www.ieeta.pt/sias