Medical image indexing and retrieval in a distributed computing environment

I. Safonau, K. Kurachka, V. Kovalev

Biomedical Image Analysis Department
United Institute of Informatics Problems, Minsk, BELARUS

What did we have

- Very large quantity of chest X-ray images
- Permanent growth of the image database
- Existing software modules: segmentation module, search module, a modified GRID software pack
- Needs for an efficient searching of similar cases in the database and some other content-based management

What did we have

Very large quantity of chest X-ray images

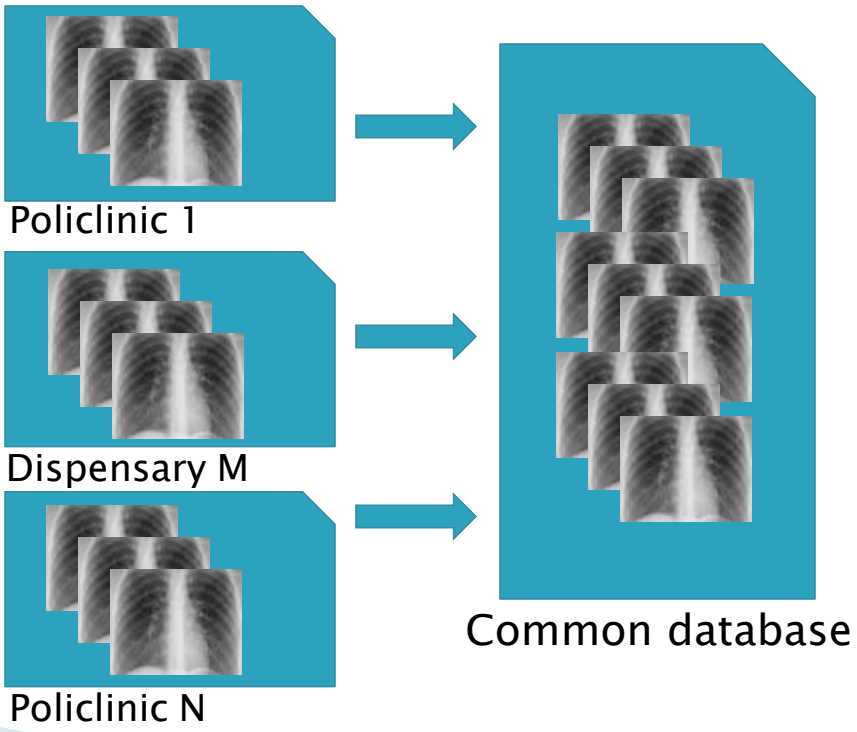
National-wide lung screening and diagnosis programme: 5 million X-ray examination per year, 900000 X-ray examination in the capital (Minsk).

Common database of X-ray images based on the data of some policlinics and TB dispensaries.

What did we have

Permanent growth of the image database

120 X-ray examination per day

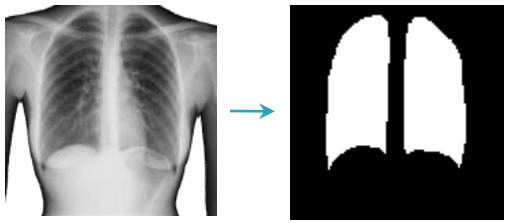


560×576 pixels,
16 bit grayscale,
650 Kb

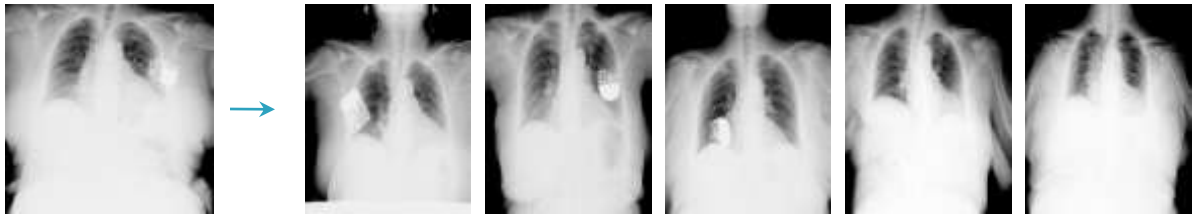
What did we have

Existing software modules

- A segmentation tool



- A searching tool (on local files)



- A modified GRID software pack (from BelGRID repository site)

Unicore/X, Gateway, XUUDB, UCC, MPICH1

What did we have

Needs for an efficient searching of similar cases in the database and some other content-based management

- Integration in existing diagnostician workstation
- Searching of similar cases
  - for learning purposes;
  - for diagnostician purposes.

What did we do

- Creation of indexing software considering applied restrictions:
  - using of certain GRID middleware;
  - existing software modules (calculating of distance measure).
- Testing of the software
  - 4 different computer configurations in local mode;
  - max. 15 compute nodes;
  - max. 4000 X-ray images (descriptors);

What did we do

Image database indexing

1. Calculating descriptors of segmented area

2. Storing the descriptors in the database for further indexation

ID	NUMBE...	NUMB...	NUMBERIS...	DAT...	DESCR	IS...	NE...	N...	MASK	MAR...	S...
51	1	1	51	06.11.2		<null>	0	0	<null>	<null>	0
52	1	1	52	06.11.2		<null>	0	0	<null>	<null>	0
53	1	1	53	06.11.2		<null>	0	0	<null>	<null>	0
54	1	1	54	06.11.2		<null>	0	0	<null>	<null>	0
55	1	1	55	06.11.2		<null>	0	0	<null>	<null>	0
56	1	1	56	06.11.2		<null>	0	0	<null>	<null>	0
57	1	1	57	06.11.2		<null>	0	0	<null>	<null>	0

What did we do

Indexing of medical image database

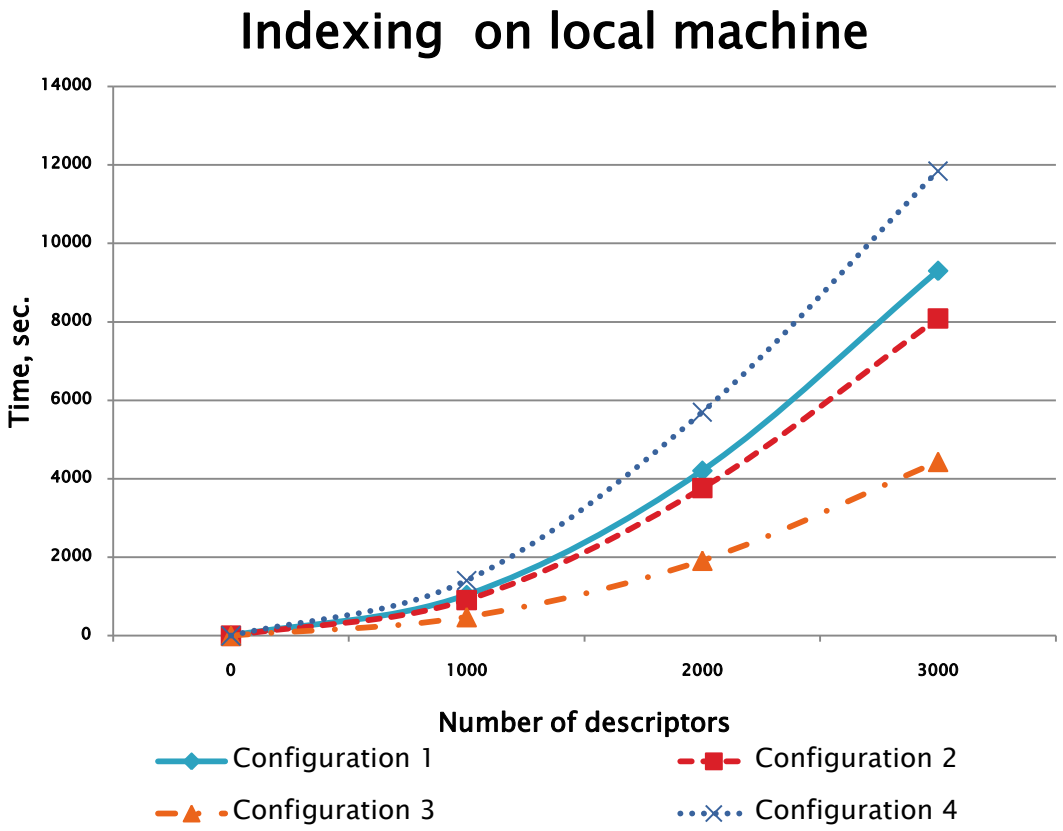
3. The database indexing

Calculating the Distance (dissimilarity) matrix for the top N most similar images

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	9377	6300	6820	2728	2265	2203	3847	5090	1789	1268	6024	5731	5709	1277	833	4009
2		0	4569	8019	8486	3846	3872	198	1488	7430	3368	7378	1107	3083	9945	1144	7156
3			0	5776	1537	9315	9461	3824	4452	2896	6051	2801	6224	5138	6587	7100	1991
4				0	4587	3811	2336	7307	1082	4485	2055	8965	2143	1673	6862	8194	4754
5					0	8689	7994	2152	1241	9374	1974	5970	8511	3554	9685	2123	2857
6						0	7412	1574	7189	346	5494	1641	5057	6657	5804	4714	6120
7							0	9025	2802	5279	5099	4878	4087	8486	8688	5211	4438
8								0	4580	2284	1045	727	8332	3845	9747	8058	2327
9									0	4931	858	9180	8408	3505	6972	1015	6979
10										0	6291	8191	6528	4658	8831	1431	550
11											0	3697	4082	7798	3276	7610	2554
12												0	560	871	7150	9384	6683
13													0	1443	6521	1422	8308
14														0	4497	5815	6221
15															0	1433	2698
16																0	3476
17																	0

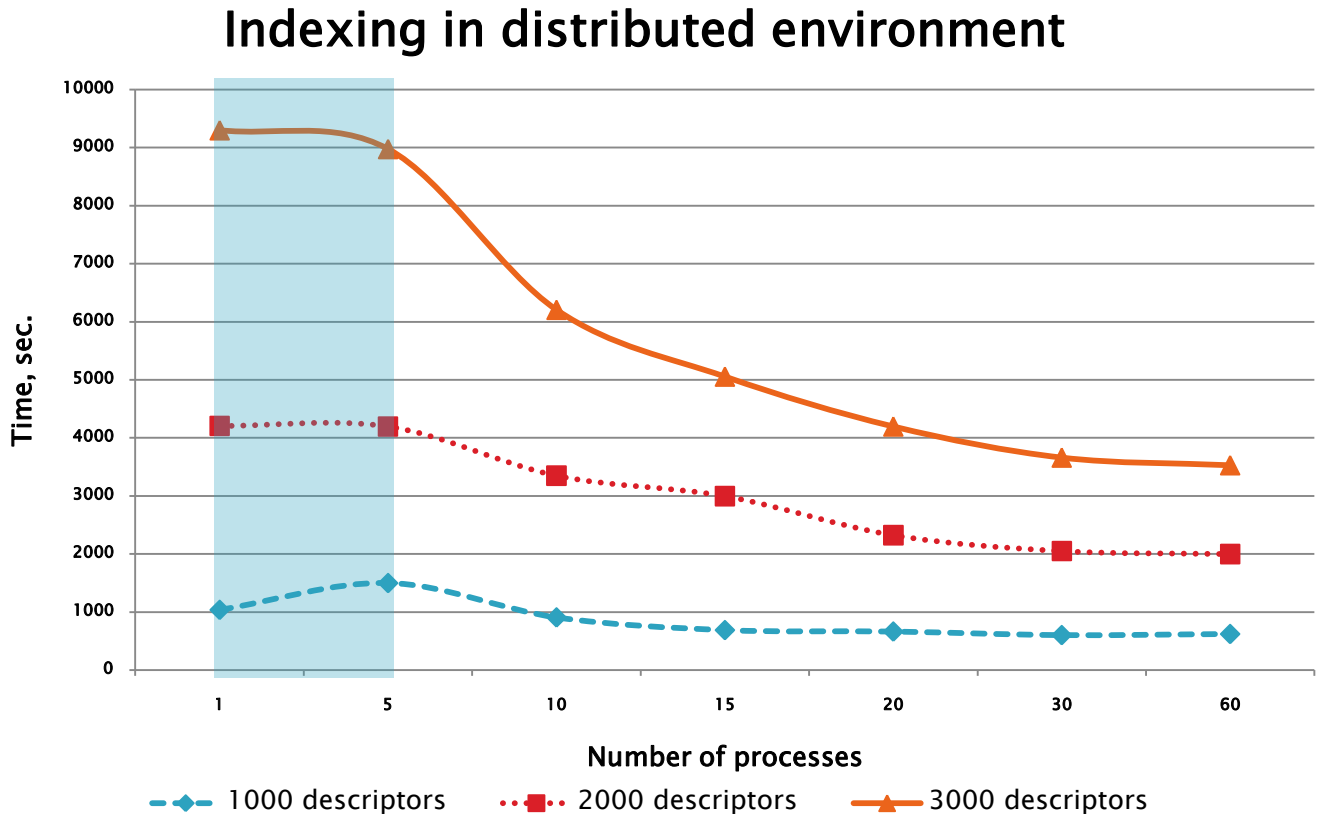
What did we do

Testing the software



What did we do

Testing of the software

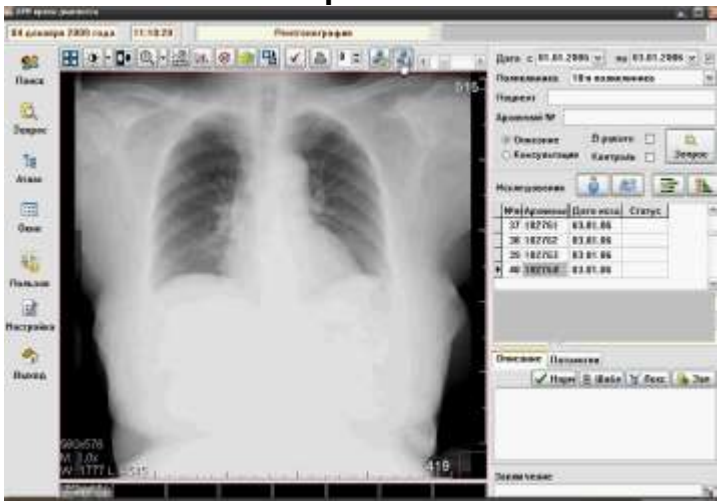


What did we do

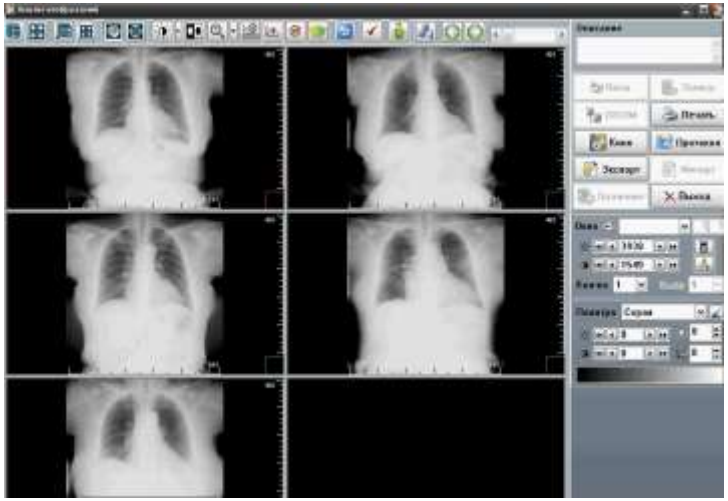
The search procedure

Integration into existing diagnostician workstation

Request



Result



Future work

- Deep testing of the software in the local network of tuberculosis dispensary (full database with more than 250000 images)
- Applying the software to a lager grid (national GRID segment, Virtual Organization)