Deep learning for all: managing and analyzing underwater and remote sensing imagery on the web using BisQue

Permalink
https://escholarship.org/uc/item/9z73t7hv

Authors
Fedorov, Dmitry, V
Kvilekval, Kristian, G
Doheny, B
et al.

Publication Date
2017-03-01

Peer reviewed
Deep learning for all
What is **BisQue**

Management, annotation and analysis of complex scientific data:

- Any metadata schema in one system
- Very large 5D images in 250+ formats (videos, 3D, ...)
- Tables, charts, chemical, genetic data, ...
- Scalable analysis
- Multi-modal searching and mining
- Web based (access anywhere with web browser)
- Cloud-based and scalable (many users and analyses)
- Commercial version and support via
Workflow

Raw Data
From any instrument: 250+ image formats
Electrophysiology
Flow Cytometry

Experimental Context
Graphical and textual descriptions
Link data between experiments

Analysis
Use your favorite tools
applied to large datasets

Visualize Results
3D+time+channel
Volume rendering

Aggregate Information
Summarization
Analytics and mining

Make a Decision
Share results
Interact remotely

UCSB system supports labs working on:
Life sciences, Materials, GIS and ecology, Medical, Marine sciences
- Multimodal: Textual, 5D graphical, images, graph, geometry
- Semantic meaning

- 5D images of any size
- WebGL volume viewer
- Many types: images, molecules, PDFs, etc...
- Plots and summaries

- Large-scale & extensible
- Automated UI
- Docker deployments
- Python, ITK, Matlab, CellProfiler, DREAM.3D, ...
BisQue | Architecture

- Connoisseur services
  - Deep learning classification/training
- Index services
  - Content indexing: Full text, graph, geometry, image content
- Query services
  - Query dialects and extensions: XPath, BICKER,...
- Blob/Image Services
- Data Services
- Analysis Services

XML, JSON, binary

- Full text
- Graph
- Geometry
- Image similarity

Python

MathWorks

Java

iRODS

Amazon web services

PostgreSQL

Spark

HTCondor
Features

BisQue | Connoisseur
Features

Dynamic organizer

Import anything

Sharing

Export everything
Features

Powerful image viewer and annotator

Volume rendering

Movies

Tabular data

Custom data types

Geographic data
Deep learning for all

BisQue | Connoisseur
Annotations by experts

Many images + Many annotators
Statistics of annotated data

Dynamic export into desired forms
Typical machine learning

Images annotations
Select features
Select classifier
Train & test

Engineer required throughout
Time consuming (3 months)
Only works on specific type of data

Best example for marine sciences

Features: Color and Texture
Classifier: SVM
Goals for Connoisseur

Automate annotations
without engineer
in reduced time
for specific datasets
Continuously improve classifier
Deep learning made easy

Benefits

Generalizes to your data
Fully automated - no feature selection
High accuracy

Leverages

Big data
Scalable services
Annotation system
Multimodal queries
Cluster processing
Fast classification on GPUs

Convolutional Neural Network (CNN)

Convolutional levels: Feature extraction typically required an engineer to select or develop a feature descriptor with CNNs it is learned from data.

Classifier: fully connected layers
Deep learning made easy

Learning pipeline:
- Images
- Annotations
- Multimodal Query Selector
- Deep Learner
- Model

BisQue services

Recognition pipeline:
- Images
- Model
- Deep Learner
- Annotations

GPU cluster
Model building

Connoisseur | Sea fruits

Dataset: Watersipora-Reef (13 images)

<table>
<thead>
<tr>
<th>Available classes</th>
<th>Model performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>185</td>
</tr>
<tr>
<td>29</td>
<td>134</td>
</tr>
<tr>
<td>143</td>
<td>111</td>
</tr>
<tr>
<td>108</td>
<td>107</td>
</tr>
<tr>
<td>103</td>
<td>98</td>
</tr>
<tr>
<td>52</td>
<td>51</td>
</tr>
<tr>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>47</td>
<td>46</td>
</tr>
<tr>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Minimum number of samples: 73
Minimum class accuracy: 20%
Minimum sample goodness: 95%

1. Select data for processing:
2. Parameters:
3. Run algorithm:
4. Results:

Connoisseur: automated image annotation
The model automatically annotates selected images.

1. Select data for processing:
- Images to classify:
- Choose a set of images
- Train/test images

2. Parameters:
- The classification method: SVM (Support Vector Machine)
- Number of support vectors:
- Kernel function:
- Parameter for the kernel function:
- Scale the data:
- Normalize the data:

3. Run algorithm:
- This may take some time.

4. Results:
The model ran in 30 seconds.

Classified image
Results
### Results

<table>
<thead>
<tr>
<th>Species</th>
<th>Samples</th>
<th>Accuracy</th>
<th>Error</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnacle (<em>Megabalanus californicus</em>)</td>
<td>460</td>
<td>87%</td>
<td>1.50%</td>
<td>0.8</td>
</tr>
<tr>
<td>Sponges</td>
<td>946</td>
<td>91%</td>
<td>2.30%</td>
<td>0.84</td>
</tr>
<tr>
<td>Anemone (<em>Anthopleura</em> sp.)</td>
<td>421</td>
<td>98%</td>
<td>0.70%</td>
<td>0.97</td>
</tr>
<tr>
<td>Echinodermata (Ophiuroid)</td>
<td>1011</td>
<td>99%</td>
<td>3.20%</td>
<td>0.94</td>
</tr>
<tr>
<td>Bryozoa (<em>Watersipora subatra</em>)</td>
<td>489</td>
<td>98%</td>
<td>1.50%</td>
<td>0.95</td>
</tr>
<tr>
<td>Anemone (<em>Metridium senile</em>)</td>
<td>4465</td>
<td>100%</td>
<td>8.20%</td>
<td>0.91</td>
</tr>
<tr>
<td>Anemone (<em>Corynactis californica</em>)</td>
<td>2496</td>
<td>100%</td>
<td>4.70%</td>
<td>0.94</td>
</tr>
<tr>
<td>No Data (water)</td>
<td>1272</td>
<td>99%</td>
<td>2.80%</td>
<td>0.95</td>
</tr>
<tr>
<td>Hydroid</td>
<td>710</td>
<td>80%</td>
<td>1.00%</td>
<td>0.76</td>
</tr>
<tr>
<td>Red Algae (filamentous)</td>
<td>1752</td>
<td>98%</td>
<td>12.10%</td>
<td>0.81</td>
</tr>
<tr>
<td>Tube Complex</td>
<td>1175</td>
<td>92%</td>
<td>2.20%</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Percent cover at 0% confidence
Percent cover at 95% confidence
Centroids
Regions (Substrate partitioning)
Semantic segmentation

Images → SLIC segmentation → Deep Learner → Goodness Filter → Segmentation
Segmentation at 0% confidence
Work in progress

Hierarchical models

- Taxonomic: Genus -> Species
- Ontological: Substrate -> Exact object

Detection workflow

Fully convolutional segmentation

Scalable deployment
BisQue | Connoisseur
Open-source BisQue (BSD)

ViQi - BisQue

- Optimized code base
- Fully licensed and GPL-free stack
- Early access to new features

ViQi SaaS service

- Scalable data and image analytics