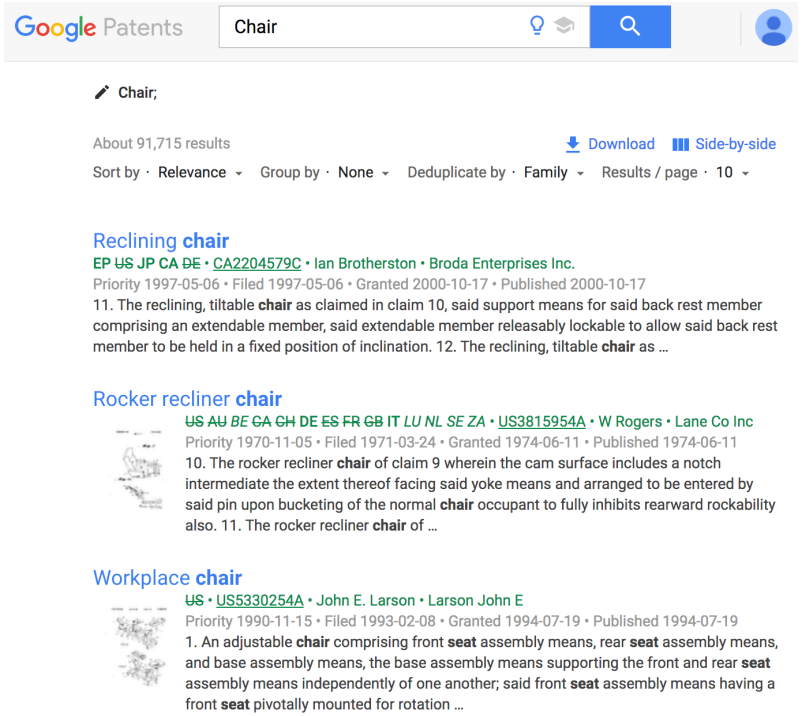
LaSPAM's
Patent Search System

Compare with Similar Services

[Google Patent]



Use common search methods for patent searches

Why we developed patent search system?

- You don’t have to do complicated search expression
- Similar patents can be identified
- You can search by patent Image
- You can be provided with a visualization service

LaSPAM's provides ···

1. Text Search

Word Cloud
Show words related to keywords searched

Landscaping & DataTable
Result of patent clustering of searched keywords

2. Image Search

Related patents
Results of patent related to image searched

Similar Image
Results of patent related to image searched

Who uses this system?

- To search for similar patents
- Who wants to see the similarity between a patent
- To find a patent ID from the patent pictures you have

Deep Patent Embedding based
Patent search & Landscaping service

Dataset

: **PatentsView** API
(supported by the US Patent and Trademark Office)

* LaSPAM used data since 2001 and type='utility'

[Data Table used for Text Search]

- **Patent** : Data concerning granted patents
- **Patent Assignee** : Crosswalk between patent and assignee tables
- **CPC current** : Current CPC classification data for all patents
- **WIPO & WIPO field** : WIPO technology classification

DataTable

Patent

Patent ID	type	country	date	title	abstract	kind
10000000	utility	US	2018.6.19	Coherent LADAR using intra-pixel quadrature detection	A frequency modulated (coherent) laser ...	B2
10000001	utility	US	2018.6.19	Injection molding machine and mold thickness control method	The injection molding machine includes a fixed platen, a moveable ...	B2
10000002	utility	US	2018.6.19	Method for manufacturing polymer film and co-extruded film	The present invention relates to: a method for manufacturing a polymer film ...	B2

Patent ID : patent this record corresponds to (Patent ID)
type : category of patent

- defensive publication : 509
- design : 613951
- plant : 26802
- reissue : 18736
- statutory invention registration : 2258
- TVPP : 3
- utility : 6426469

country : country in which patent was granted (always US)
date : date when patent was granted
title : title of patent
abstract : abstract text of patent
kind : WIPO document kind codes

Use since 2001 and type = 'utility' **4,201,248** rows of 5,992,136 rows

DataTable

Patent Assignee

patent ID	assignee ID
10000000	341194
10000001	133905
10000002	385861

Patent ID : patent_number

Assignee ID : unique assignee id generated (by the disambiguation algorithm)

CPC Current

Patent ID	Section ID	Subsection ID	Group ID	Subgroup ID	category
10000000	G	G01	G01S	G01S7/4863	inventional
10000000	G	G01	G01S	G01S7/4865	inventional
10000000	G	G01	G01S	G01S7/4914	inventional

Patent Id : patent number

Section ID : CPC section

Subsection ID : CPC subsection ID

Group ID : group ID

Subgroup ID : subgroup ID

category : CPC category

DataTable

WIPO & WIPO Field - WIPO technology classification

Patent ID	Field ID	Sector title	Field title
10000038	1	Electrical engineering	Electrical machinery, apparatus, energy
10000058	1	Electrical engineering	Electrical machinery, apparatus, energy
10000097	1	Electrical engineering	Electrical machinery, apparatus, energy

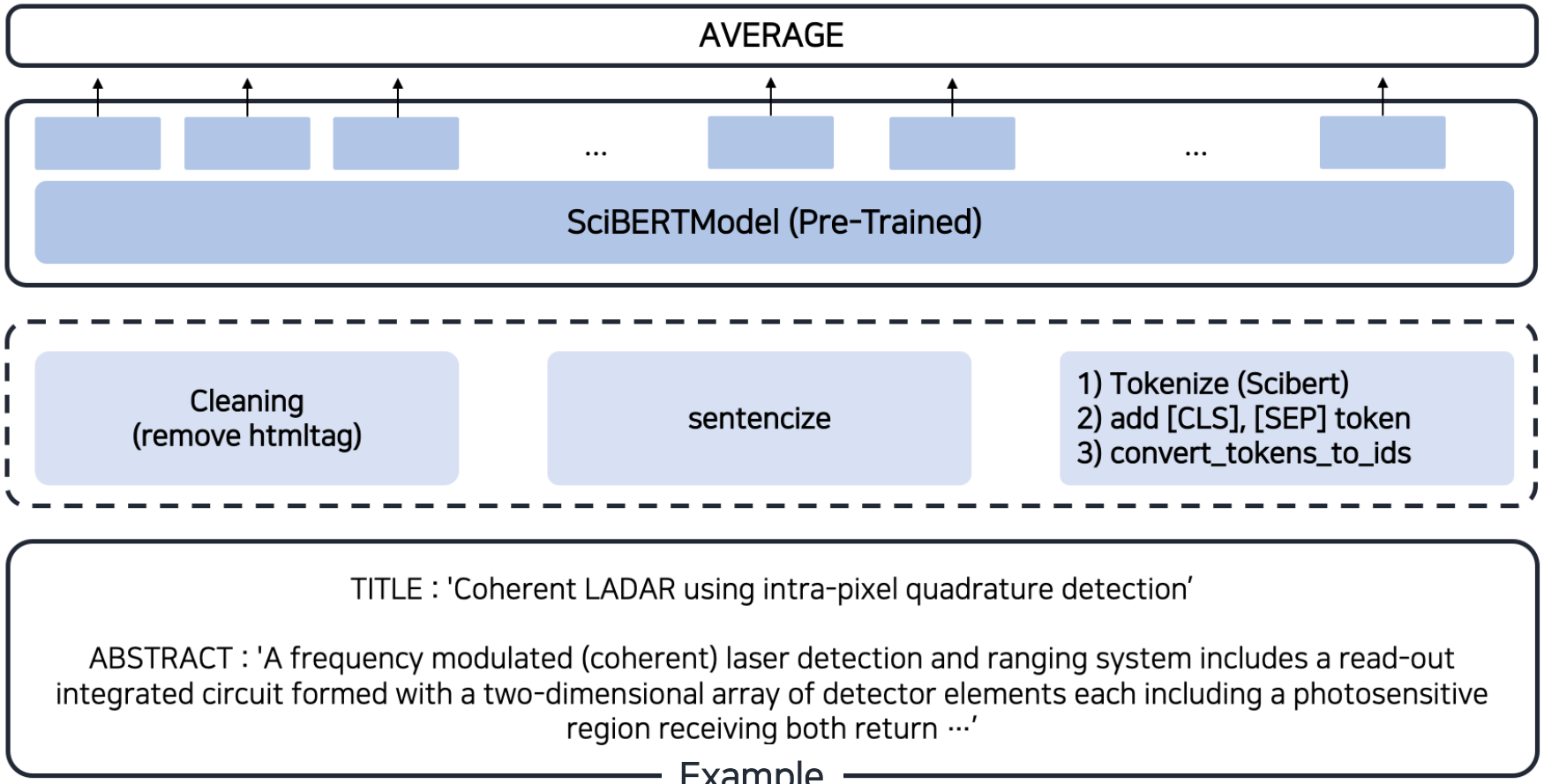
Patent ID : patent number

Field ID : WIPO technology field id as derived from crosswalk

Sector title : WIPO technology sector title

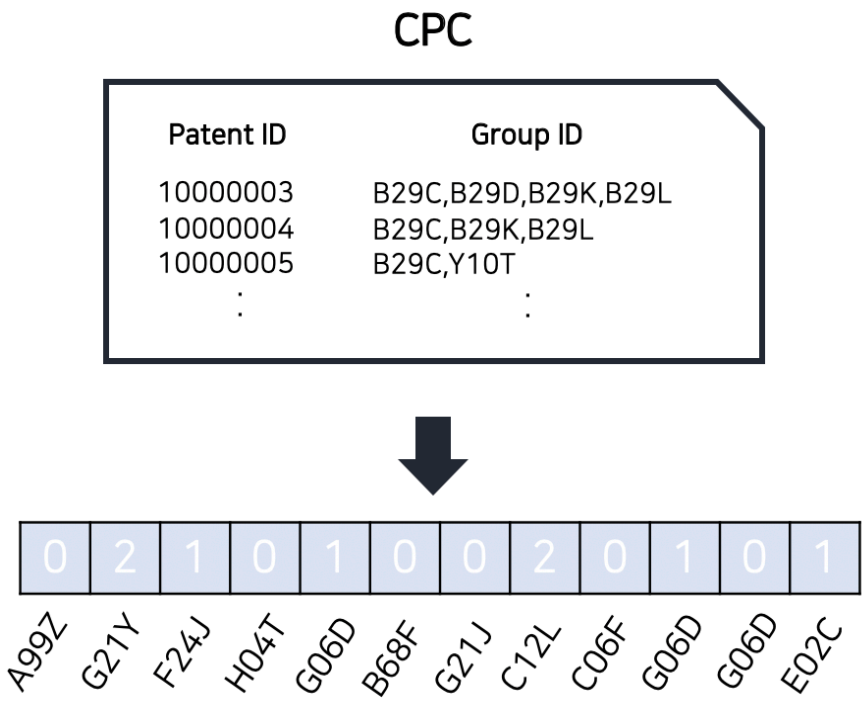
Field title : WIPO technology field title

Text Preprocess
for Patent Embedding Model



Metadata Preprocess
for Patent Embedding Model

- CPC : bag of words
  - Features : Patent ID, Group ID
- Assignee : bag of words
  - Features : Patent ID, Assignee ID
- WLPO : bag of words
  - Features : Patent ID, WIPO Sector Title
  - Apply argmax for categorical classification

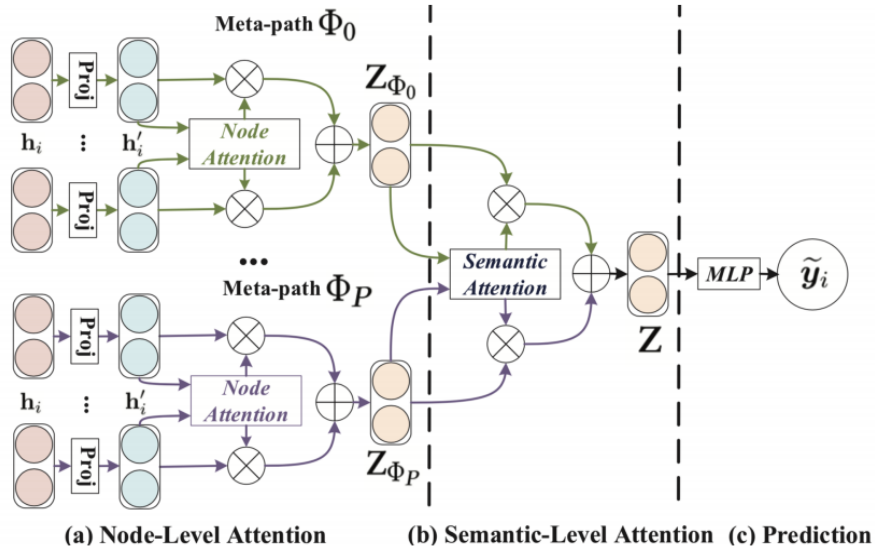


[Example]

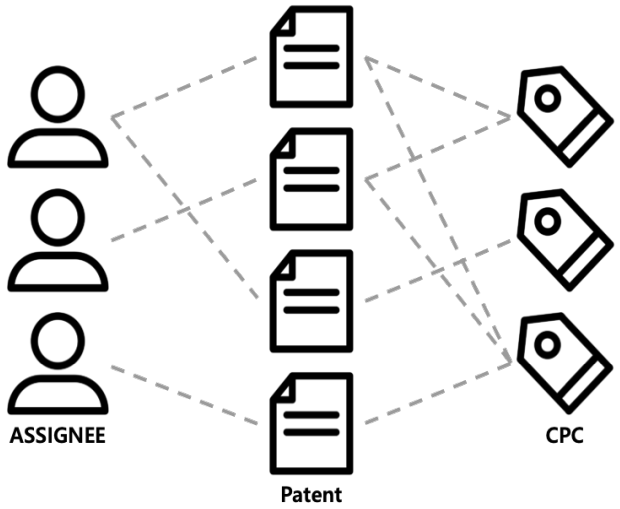
Patent Embedding Model

What algorithm is used?

HAN (heterogeneous attention network)



[Heterogeneous Graph for Patent]



- Deep Graph Library (DGL)
- Pytorch

Model Results

- Train Micro f1 : 0.8963
- Train Macro f1 : 0.8313

Deep Image Representation Vector based Image search service

Dataset

: **PatentsView** API
(supported by the US Patent and Trademark Office)

* LaSPAM used data since 2001 and type='utility'

[Data Table used for Image Search]

- **Patent** : Data concerning granted patents
- **CPC current** : Current CPC classification data for all patents

\+ crawled thumbnail **image data** of Patent ID

ex) **Images (5)**

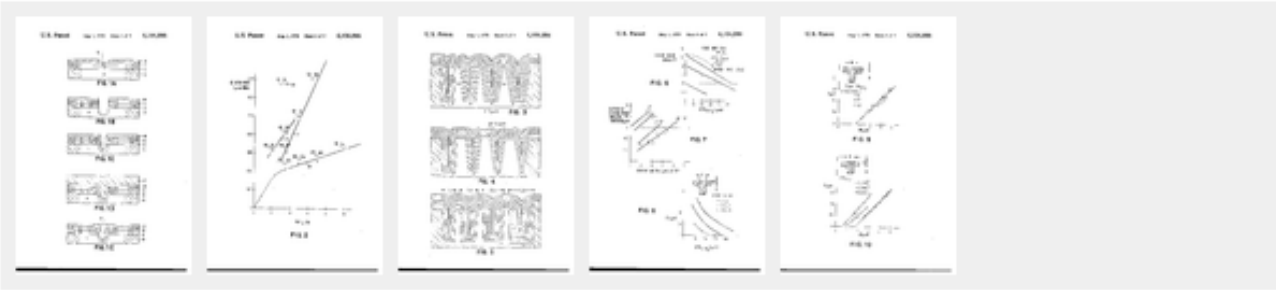
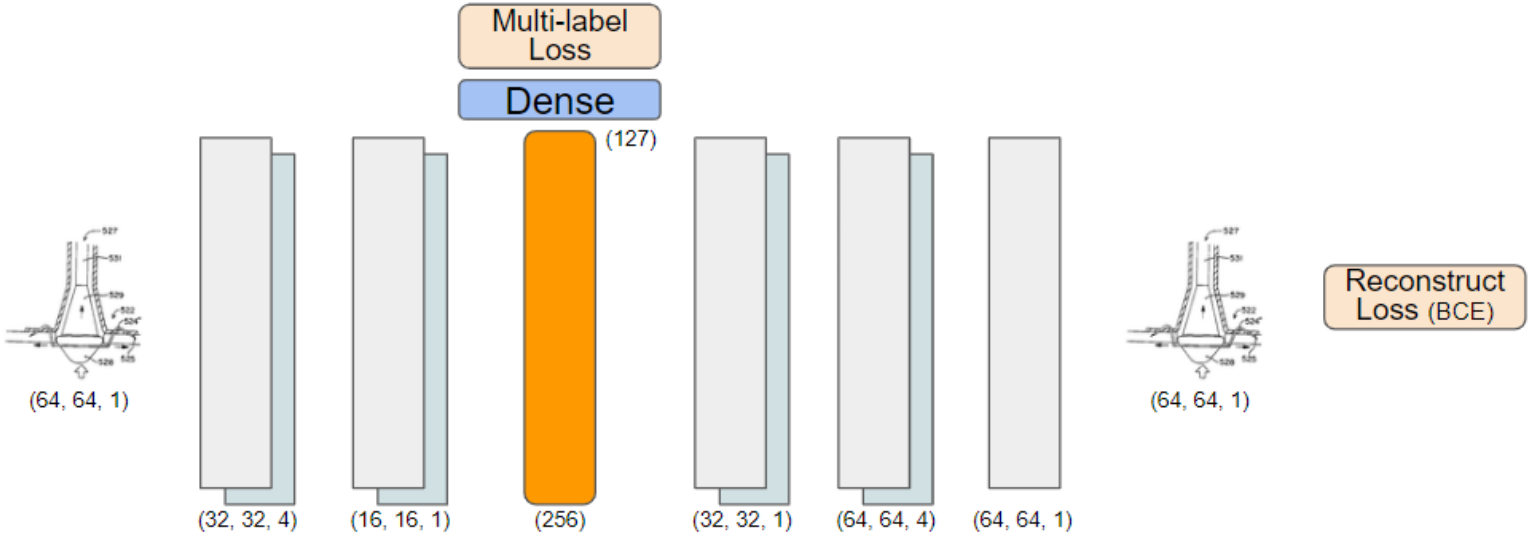


Image Search Algorithm - Model

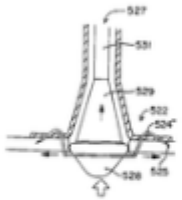
Convolutional – Deconvolutional Autoencoder



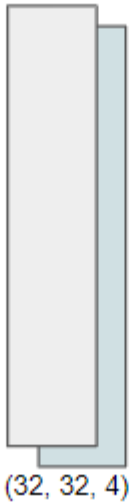
- Resize to 64 * 64 images crawled from Google patents thumbnail
- Encoder consist of conv – maxpool and Decoder consist of deconv – upsample
- Embedding layer occurs 3-digit cpc code(127 labels)
  classification loss with Dense - Sigmoid Layer
  (To induce embedding values to have informations about the cpc code.)

Image Search Algorithm - Searching

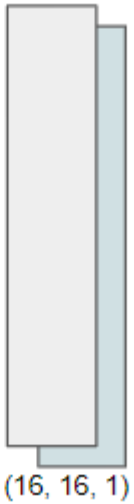
Convolutional – Deconvolutional Autoencoder



Input image



(32, 32, 4)



(16, 16, 1)

(256)

Cosine similarity

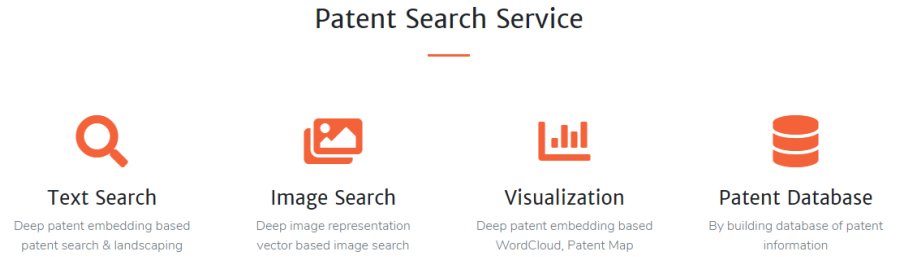
0.947	0.548	0.248	0.775	···	0.159
0.157	0.489	0.684	0.442	···	0.638
		···			
		···			
		···			
0.613	0.348	0.745	0.145	···	0.575

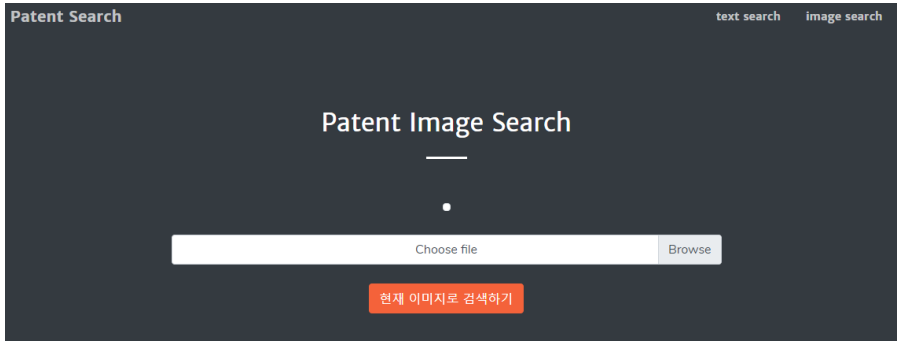
- CPC code classification(Multi-label classification) Precision : 0.116
- Top-5 accuracy with searched patents : 0.99

LaSPAM's
Patent Search System

LaSPAM's Patent Search

http://laspam.com



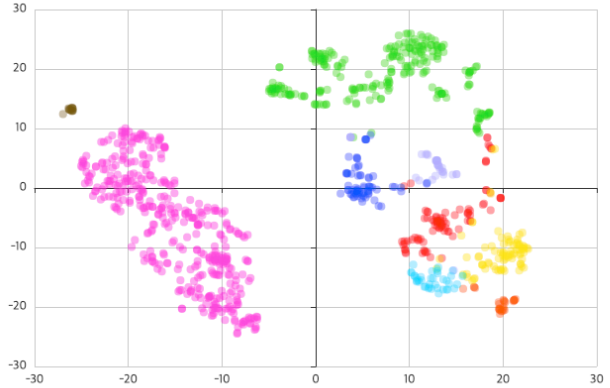


LaSPAM's Patent Search

: Patents Text Search
Deep Embedding based
Patent Search & Landscaping

[Demonstration]

[Patent Landscape]

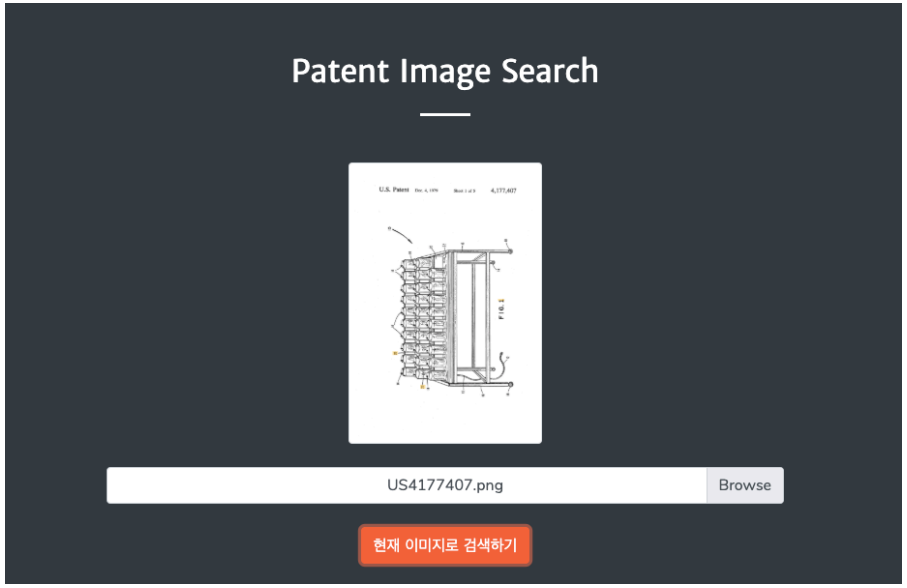


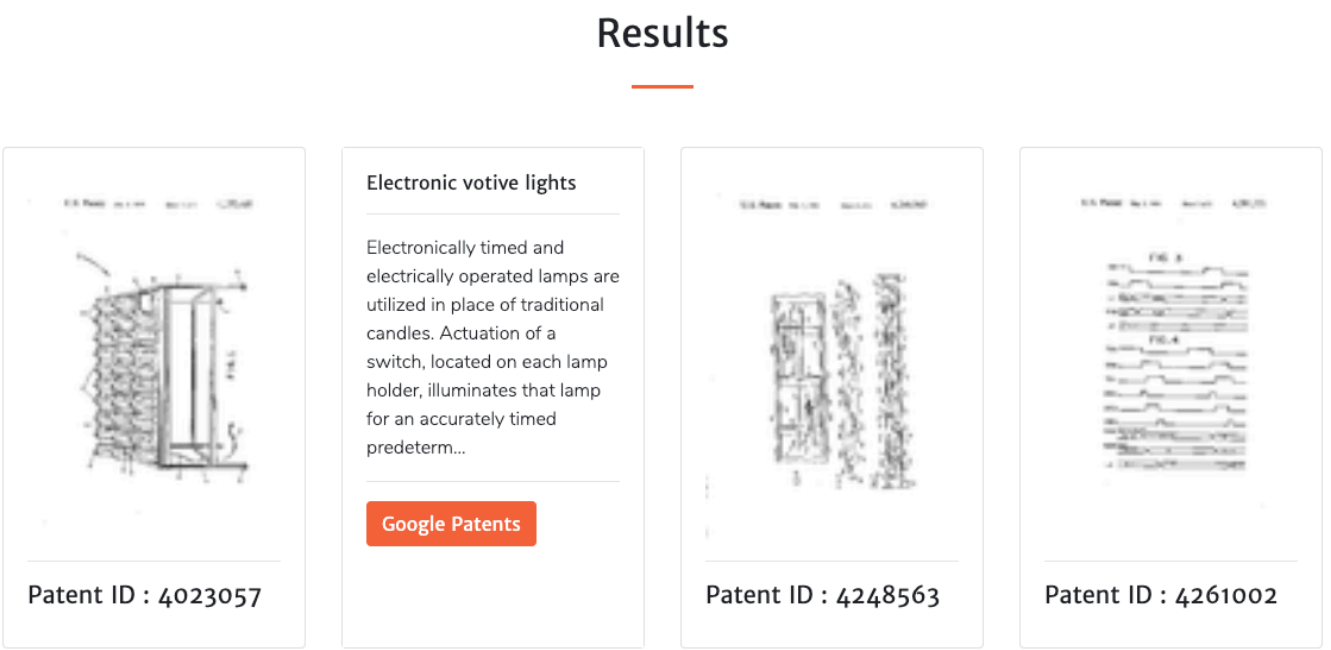
[WordCloud]

LaSPAM's Patent Search

: Patents Image Search
Deep Image Representation
Vector based Image Search

[Demonstration]





Thank you