

Movie Poster Images Dataset

1. Introduction

This dataset was collected from the IMDB website. One poster image was collected from one (mostly) Hollywood movie released from 1980 to 2015. Each poster image is associated with a movie as well as some metadata like ID, genres, and box office. The ID of each image is set as its file name.

2. Data Description

Movie_Poster_Dataset.zip – The poster images

Movie_Poster_Metadata.zip – Metadata of each poster image

For each poster, the associated metadata are represented as follows.

```
{
  "_id" : ObjectId("56fb860e71a0651b0c306f55"),
  "Language" : "English",
  "Box_office" : "290,158,751",
  "Country" : "USA",
  "Rated" : "PG",
  "imdbID" : "tt0080684",
  "Awards" : "Won 1 Oscar. Another 15 wins & 18 nominations.",
  "Poster" : "http://ia.media-imdb.com/images/M/MV5BMjE2MzQwMTQxN15BM15BanBnXkFtZTcwMDQzNjk2OQ@@.V1_SX300.jpg",
  "Director" : "Irvin Kershner",
  "Released" : "20 Jun 1980",
  "Writer" : "Leigh Brackett (screenplay), Lawrence Kasdan (screenplay), George Lucas (story)",
  "imdbVotes" : "799,579",
  "Runtime" : "124 min",
  "Response" : "True",
  "imdbRating" : "8.8",
  "Title" : "Star Wars: Episode V - The Empire Strikes Back",
  "Genre" : "Action, Adventure, Fantasy",
  "Actors" : "Mark Hamill, Harrison Ford, Carrie Fisher, Billy Dee Williams",
  "Type" : "movie",
  "Metascore" : "79"
}
```

Meanings of metadata:

“**_id**” means the object ID of this movie on IMDB.

“**Language**” means the spoken language of this movie.

“**Box_office**” means the total box office of this movie.

“**Country**” means where the movie was released.

“**Rated**” means the rate of this movie, for example, “PG”, “R”.

“**imdbID**” means the ID of this movie from IMDB. This is the main ID we indicate a poster image. You can find the corresponding movie poster image with this ID.

“**Awards**” means the awards of this movie.

“**Posters**” means the original URL of this movie poster.

“**Director**” means the director of this movie.

“**Released**” means when this movie was released.

“**Writer**” means the writer of this movie.

“imdbVotes” means the votes of this movie on the IMDB website.

“Runtime” means the video length of this movie.

“Response” means whether there are someone response this movie or not.

“imdbRating” means the rating of this movie on the IMDB.

“Title” means the title of this movie.

“Genre” means the genres of this movie.

“Actors” means all the actors of this movie.

“Type” means the movie type, for example, “movie”, “series”

“Metascore” means the metascore of this movie.

3. Citation

If you use this dataset, please refer and cite our work published in the following:

Wei-Ta Chu and Hung-Jui Guo, “Movie Genre Classification based on Poster Images with Deep Neural Networks,” Proceedings of International Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes, pp. 39-45, 2017. (in conjunction with ACM Multimedia 2017)