

A Feature-Enhanced Ranking-Based Classifier for Multimodal Data and Heterogeneous Information Networks

Scott Deeann Chen, Ying-Yu Chen, Jiawei Han, Pierre Moulin

University of Illinois at Urbana-Champaign

Dec 8, 2013

Overview

Feature-enhanced RankClass (F-RankClass):

- ▶ Classification algorithm on heterogeneous information networks
- ▶ Both unimodal and multimodal data
- ▶ Extends naturally from binary to multiclass problems

Multimodal document classification

Experimental results

Multimodal Document Classification

Multimodal Documents

- ▶ Multiple modalities: text, image, audio, video

Campus [edit]

Main article: [Campus of the University of Illinois at Urbana–Champaign](#)

The campus is known for its landscape and architecture, as well as distinctive landmarks.^[23] It was identified as one of 50 college or university 'works of art' by T.A. Gaines in his book *The Campus as a Work of Art*.^[24]

The main research and academic facilities are divided almost evenly between the twin cities of [Urbana](#) and [Champaign](#). The [College of Agriculture, Consumer, and Environmental Sciences'](#) research fields stretch south from Urbana and Champaign into Savoy and Champaign County. The university maintains formal gardens and a conference center in nearby Monticello at Allerton Park.

U of I is one of the few educational institutions to own an airport.^[25] [Willard Airport](#), named for former University of Illinois president Arthur Cutts Willard, is located in Savoy. It was completed in 1945 and began service in 1954. Willard Airport is home to University research projects and the University's [Institute of Aviation](#), along with flights from American Airlines.

The campus is based on the [quadrangle](#) design popular at many universities. Four main quads compose the center of the university and are arranged from north to south. The Beckman Quadrangle and the [John Bardeen Quadrangle](#) occupy the center of the [Engineering Campus](#). [Boneyard Creek](#) flows through the John Bardeen Quadrangle, paralleling [Green Street](#). The Beckman Quadrangle is primarily composed of research units and laboratories, and features a large solar calendar consisting of an obelisk and several copper fountains. The Main Quadrangle and South Quadrangle follow immediately after the John Bardeen Quad. The former makes up a large part of the [Liberal Arts and Sciences](#) portion of the campus, while the latter comprises many of the buildings of the [College of ACES](#) spread across the campus map.^[26]



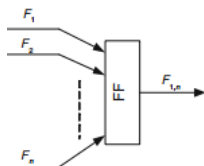
A panorama facing north on UIUC's Main Quad



Motivation: Multimodal Fusion [Atrey 2010]

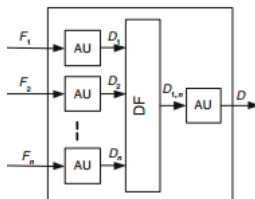
Early fusion: concatenate features and analyze upon

- ▶ structure of information not considered



Late fusion: analyze each modality independently and fuse results

- ▶ cross-modality correlation not considered



Motivation: Multimodal Document Analysis

Some media objects are only weakly correlated or even unrelated



Motivation: Multimodal Document Analysis



Which category does this image come from?

- ▶ History
- ▶ People
- ▶ Information technology
- ▶ Sports
- ▶ Politics

Motivation: Multimodal Document Analysis



Figure : Bjarne Stroustrup

Motivation: Multimodal Document Analysis

C++

From Wikipedia, the free encyclopedia



This article **may be too technical** for most readers to understand.

C++ (pronounced *see plus plus*) is a [programming language](#) that is general purpose, [statically typed](#), [free-form](#), [multi-paradigm](#). C++ was originally named **C with Classes**, adding [object oriented](#) features, such as [classes](#), and other enhancements.

C++ is one of the most popular programming languages^{[5][6]} and is implemented on a wide variety of hardware and operating system architectures. It is used for [client applications](#), and entertainment software such as [video games](#).^[7] Several groups provide both free and proprietary implementations.

The language began as enhancements to [C](#), first adding [classes](#), then [virtual functions](#), [operator overloading](#), [multiple inheritance](#). The language was amended by the 2003 technical [corrigendum](#), *ISO/IEC 14882:2003*. The current standard extending C++ with [modules](#) is *ISO/IEC 14882:2020*.

Contents [\[show\]](#)

History



Bjarne Stroustrup, creator of C++

Bjarne Stroustrup, a Danish and British trained computer scientist, began working on C++ in 1979. He was very helpful for large software development, but the language was not very useful for [kernel](#) with respect to [distributed computing](#). Remembering his Ph.D. experience with [FORTRAN](#), [Pascal](#), and [C](#), he looked for programming languages that inspired him were [ALGOL 68](#), [Ada](#), [CLU](#) and [ML](#). At first, he called the language *C with Classes*.

In 1983, the name of the language was changed from *C with Classes* to *C++*. The language was improved with [type checking](#), and [BCPL](#) style [single-line comments](#) with two forward slashes. The first commercial implementation was [C++](#) language, as there was not yet an official standard.^[11] The first commercial implementation included [templates](#), [exceptions](#), [namespaces](#), [new casts](#), and a [Boolean](#) type.

As the C++ language evolved, the standard library evolved with it. The first standard library, was a large amount of the [Standard Template Library](#).

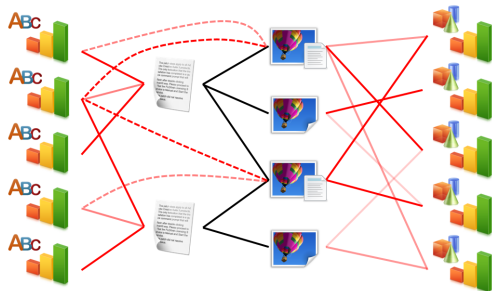
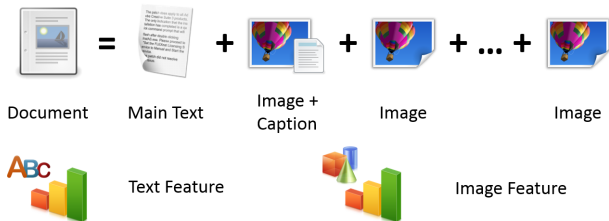
Feature-Enhanced RankClass

F-RankClass

- ▶ Based on RankClass [Ji KDD'11]
- ▶ Heterogeneous information network: nodes representing different types of objects.
- ▶ Every media object and **feature** is a node.
- ▶ Edges connect related media objects (co-existing) .
- ▶ Edge weights represent feature values.
- ▶ **Ranking** identifies typical and important nodes
- ▶ **Classification** makes ranking more meaningful



F-RankClass for Multimodal Documents



Datasets

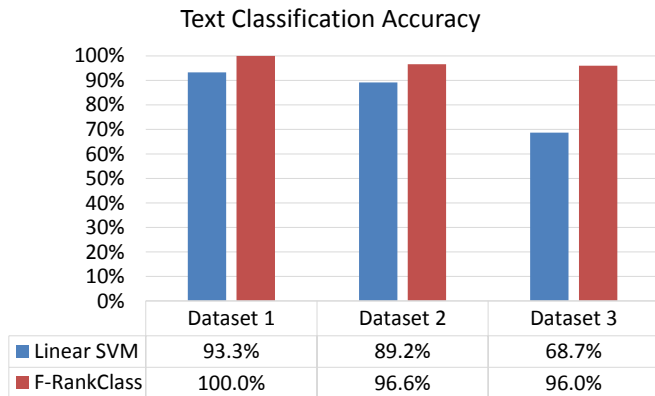


WIKIPEDIA
The Free Encyclopedia

2008/9 Wikipedia Selection for Schools

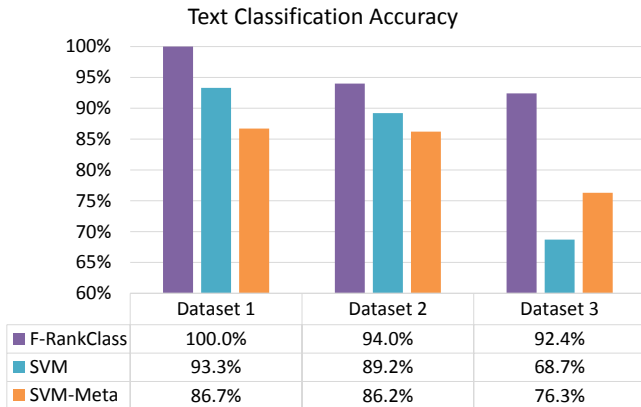
- ▶ **Real-world** dataset, **user-generated** contents, **publicly** available
- ▶ High-quality documents with related text and images
- ▶ 15 categories, 5,500 articles, 34,000 images
- ▶ Selected subsets: (1) Art, IT (2) History, Science (3) Business, Language/Literature, Math, Religion
- ▶ <http://www.ifp.illinois.edu/~chen124/wikipedia/>

Unimodal Text Classification

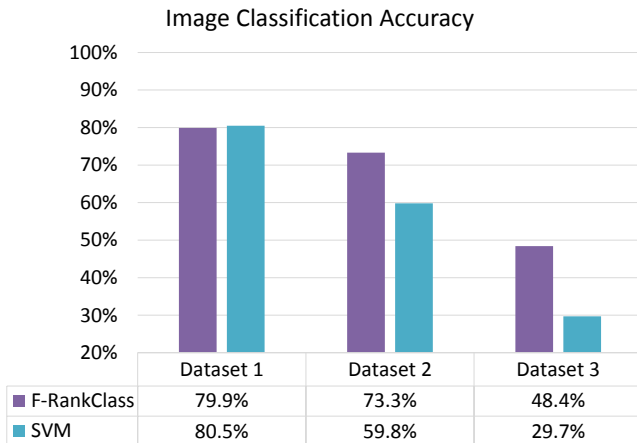


Multimodal Document Classification: Text Accuracy

Compared to SVM-Meta (late fusion), F-RankClass does not suffer from weakly correlated images.



Multimodal Document Classification: Image Accuracy



Conclusion

- ▶ Feature-enhanced RankClass (F-RankClass):
 - ▶ Classification algorithm
 - ▶ Both unimodal and multimodal data
 - ▶ Extends naturally from binary to multiclass problems
- ▶ The use of Wikipeda for Schools 2008/9 for multimodal document classification.