FITCHBURG STATE UNIVERSITY

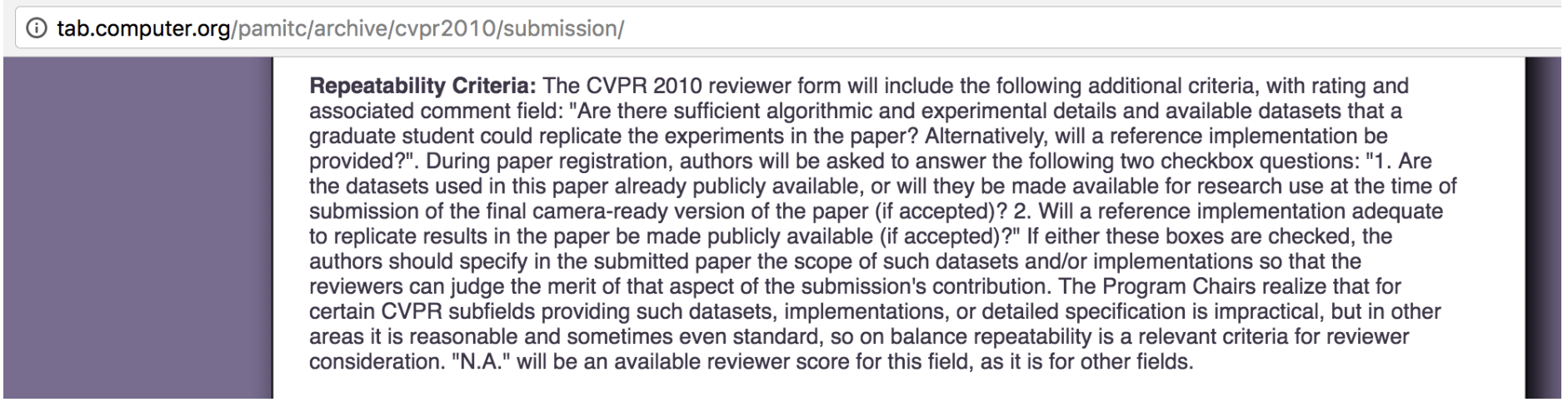
USC University of Southern California
Information Sciences Institute

REPRODUCIBILITY IN COMPUTER VISION: TOWARDS OPEN PUBLICATION OF IMAGE ANALYSIS EXPERIMENTS AS SEMANTIC WORKFLOWS

Ricky J. Sethi (FSU) and Yolanda Gil (USC/ISI)
Presented by Daniel Garijo (USC/ISI). eScience 2016

Reproducibility in Computer Vision

- The importance of reproducible computational research has come to the forefront in computer vision
- Premier conferences like Computer Vision and Pattern Recognition (CVPR) requiring reviewers to comment on the reproducibility of papers
  - The International Conference on Image Processing (ICIP) has round tables on reproducibility



Overview

- Reproducibility Crisis
- Addressing reproducibility with scientific workflows
- Case Study: Video Activity Recognition
- Case Study: Multimedia Analysis
- Case Study: Neural Algorithm of Artistic Style
- Benefits of scientific workflows for computer vision analysis
- Conclusions

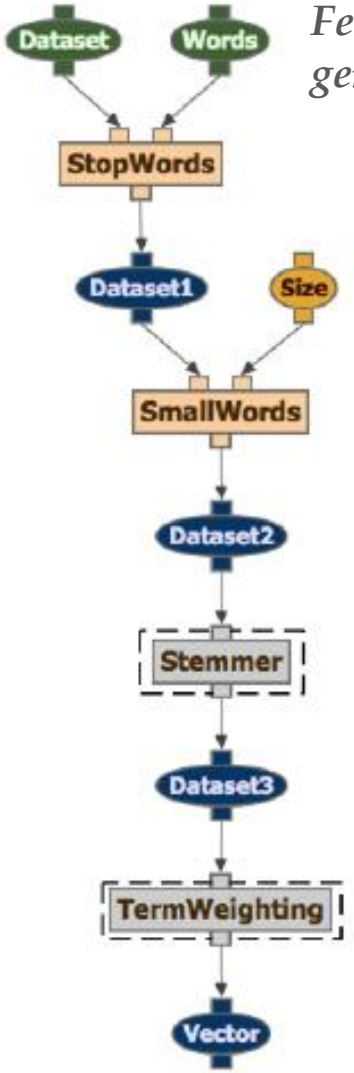
Addressing reproducibility with scientific workflows…

- General technique for describing and enacting a process
- Capture complex analytical processes at various levels of abstraction
  - Visually describes what you want to do
- Tracks metadata, parameters, and intermediate results
  - Debugging, inspectability
- Accommodate large amounts of data and large number of computations
- Semantic Workflows incorporate semantic constraints about datasets and workflow components
  - Used to create and validate workflows and to generate metadata for new data products

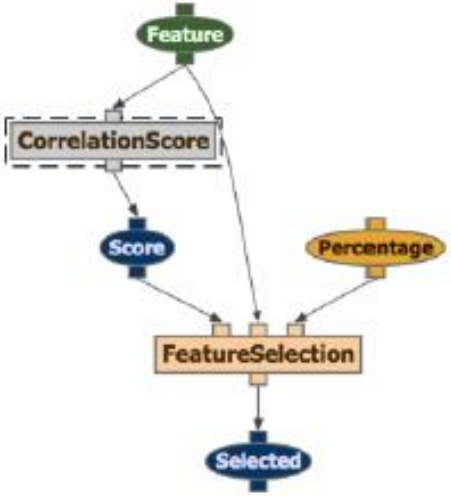
Examples of Scientific Workflows

Workflows from [Hauder, et al., SC WORKS 2011]

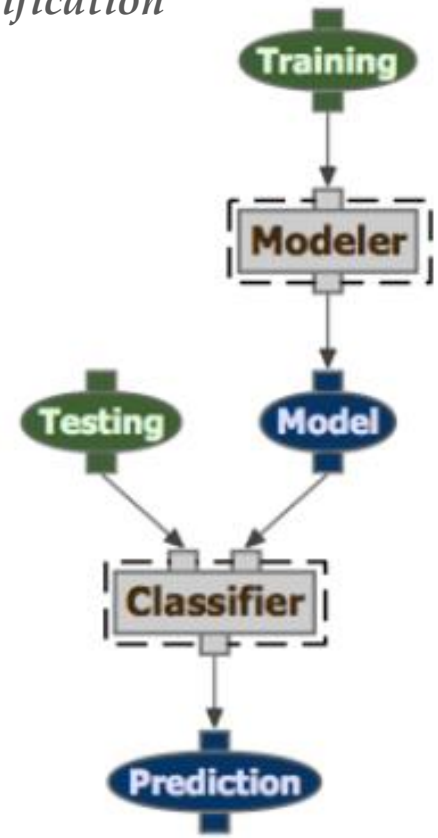
Feature generation



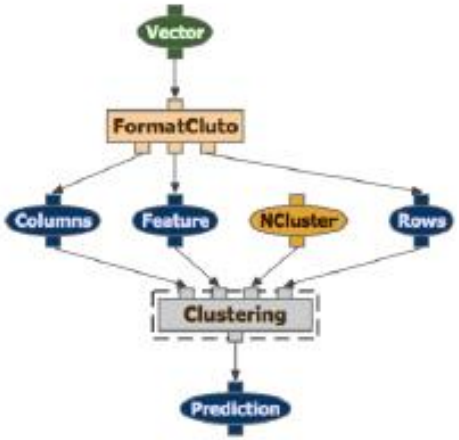
Feature selection



Classification



Clustering



Creating workflows: WINGS

- WINGS is a semantic workflow system that assists scientists with the design of computational experiments.
- Workflow representations incorporate **semantic constraints** about datasets and workflow components, and are used to create and validate workflows and to generate metadata for new data products.
- WINGS submits workflows to execution frameworks such as Pegasus and OODT to run workflows at large scale in distributed resources.

http://wings-workflows.org/

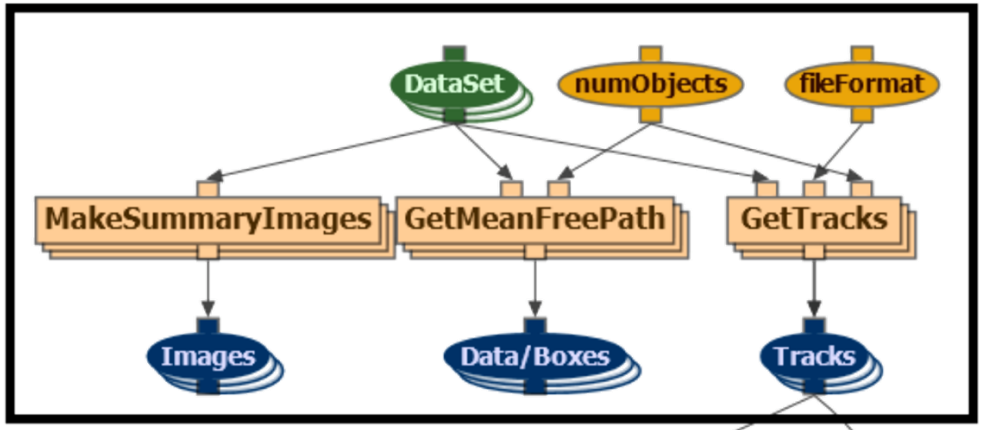
Overview

- Reproducibility Crisis
- Addressing reproducibility with scientific workflows
- Case Study: Video Activity Recognition
- Case Study: Multimedia Analysis
- Case Study: Neural Algorithm of Artistic Style
- Benefits of scientific workflows for computer vision analysis
- Conclusions

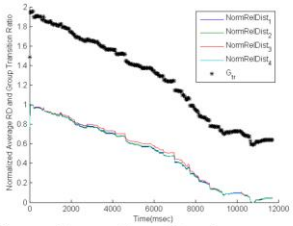
Case Study: Detecting Groups in Videos

- How can we figure out when we go from a collection of individuals to formation of a crowd in video?
- Reminiscent of the `n`-body problem in fluid dynamics: the transition from a collection of individual particles to a fluid

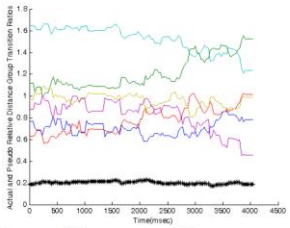
Workflows for Group Analysis

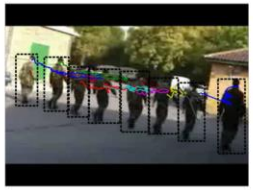


Workflow fragment for GTR



(a) *Group Formation*: G_{tr} decreases from about 2 to about 0.5. Sample image shows individuals converging and forming a group.





Workflow fragment for Statistical Evaluation

(b) *Group Movement*: G_{tr} stays steady at about 0.2. Sample image shows group moving together.

Computer Vision Workflows

- Workflow Fragments created for Computer Vision

Category	Workflow Fragments
Computer Vision	OpenCV components (Optical Flow, Kalman tracker, Mixture of Gaussians, Particle filter, etc.), N-Cuts, PhaseSpace, G_{tr}, $Relative_{Velocity}$, $Relative_{Distance}$, Image Extractor, Background/Foreground Extraction, Neural Algorithm for Artistic Style (Lua/Torch), Neural Algorithm for Artistic Style (TensorFlow)
General Machine Learning	K-Means, Latent Dirichlet Allocation, Mallet, libSVM, Caffe, Convolutional Neural Networks (Lua/Torch), TensorFlow, Adam Optimizer, Recurrent Neural Networks
Statistical Evaluation	Confusion Matrices, Heatmaps, Precision-Recall Curves, ROC Curves, AUC Curves, Equal Error Rate, F-Measure

Overview

- Reproducibility Crisis
- Addressing reproducibility with scientific workflows
- Case Study: Video Activity Recognition
- Case Study: Multimedia Analysis
- Case Study: Neural Algorithm of Artistic Style
- Benefits of scientific workflows for computer vision analysis
- Conclusions

FITCHBURG STATE UNIVERSITY

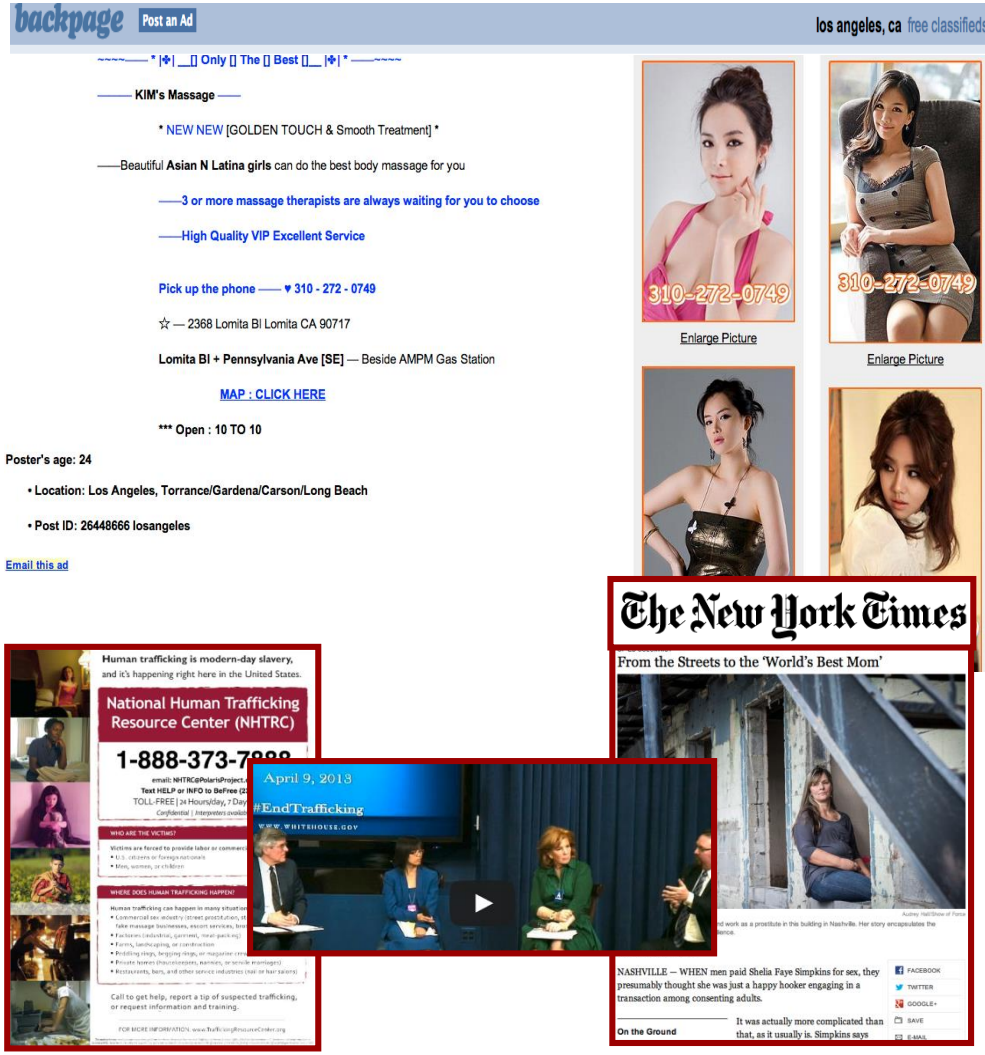
USC University of Southern California
Information Sciences Institute

Motivation: Human Trafficking Detection

- **2M children** estimated to be exploited by the global trafficking trade

- 12.3M individuals worldwide as forced laborers, bonded laborers or trafficking victims. **1.39M of them worked as trafficked slaves, 98% are women and girls**

- Global profits estimated to be **US$ 31.6B from trafficked victims,** from forced laborers US$ 44.3B per year. The largest profits - more than US$ 15B - are in industrialized countries



The Need for Automation of Human Trafficking Detection

Law enforcement activities such as tracking and capture (sting) operations are more effective through monitoring on-line ads across sites

AD CHARACTERISTICS

- Falsifying information
  - E.g. age
- Obscuring information
- Use of aliases
- Across locations

Currently done by hand!

TASKS

- Extract service modality, detect illicit services
- Estimate true age
- Link ads of same provider
- Link ads across sites/locations
- Cross-reference with DBs (e.g., missing children)

FITCHBURG STATE UNIVERSITY

USC University of Southern California
Information Sciences Institute

Multimedia Analysis for Human Trafficking Detection

IMAGE ANALYSIS

- Image age estimation/age projection
- Match face with likely victims (e.g., runaways/abductees)
- Detect multiple faces; co-trafficking highly correlated with underage participation
- Use of stock/photoshopped images inversely correlated with underage participation
- Reuse of banner images may indicate association/sharing
- ID/matching of locations (hotel decor), personal effects, tattoos even if face has been obscured
- Race/ethnicity/body characteristics estimation

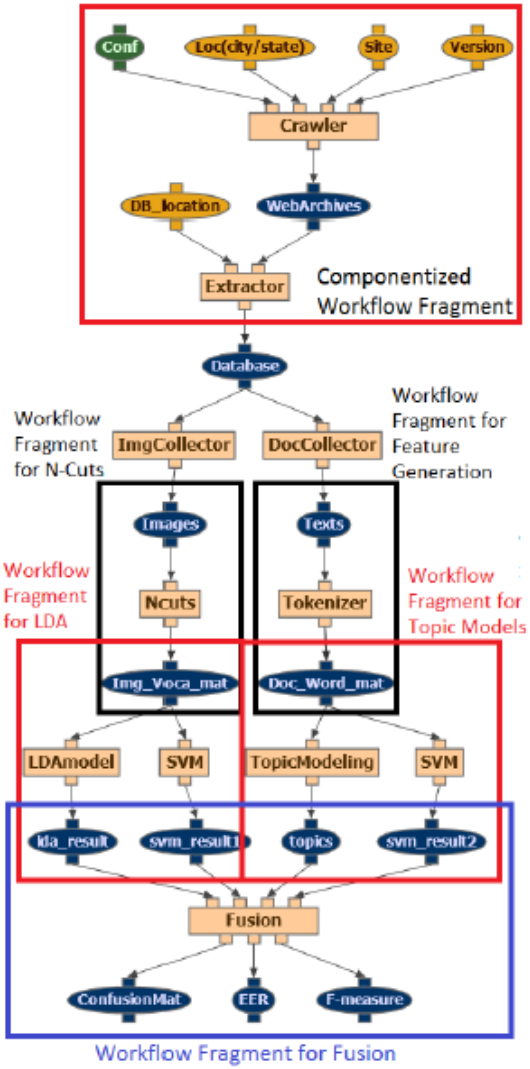
TEXT ANALYSIS

- Text indications of underage participation ("young") weaker than other methods; very often deceptive/false
- Text indication of race/ethnicity/body also have high degree of deception
- Text descriptions of co-trafficking (multiple victims) have been found to be more reliable

Combining text and image cues narrows search more effectively

TrafficBot project: 6 sites, each 400 locations, 20,000-40,000 posts/day

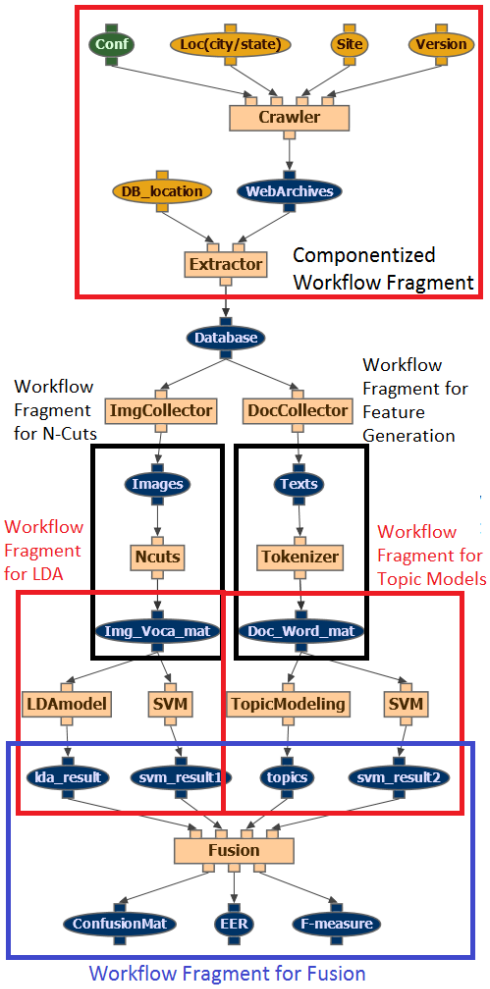
High-Level Workflow for Multimedia Analysis



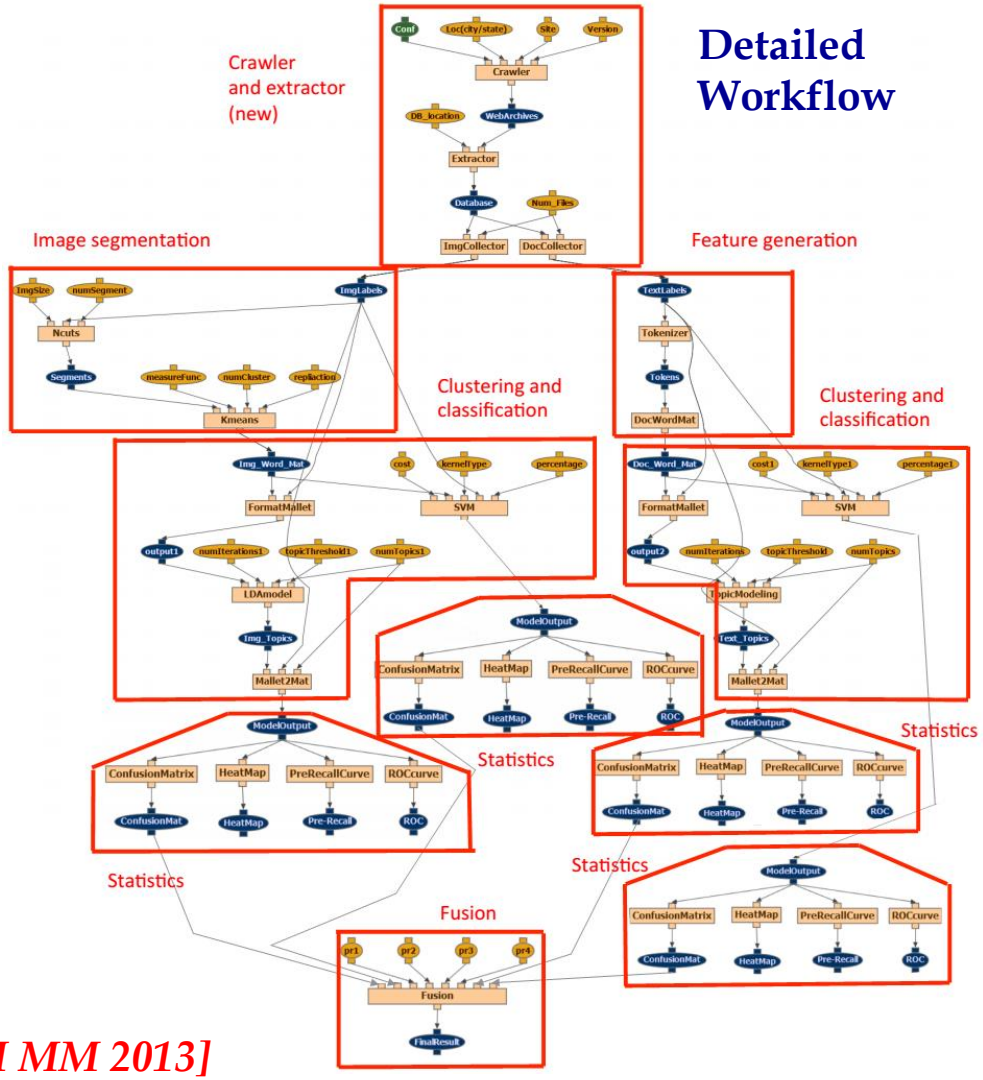
- Workflow shows the following modules:
  - **Componentized Workflow Fragment**
  - **N-Cut segmentation** on the image
  - **Workflow Fragment for Feature Generation,** as well as doing feature selection
  - **Workflow Fragment for Fusion:** combines the results from the Image Analysis (LDA and SVM) as well as the results from the Text Analysis (Topic Models and SVM).

Workflow for Multimedia Analysis

High-Level Workflow



Detailed Workflow

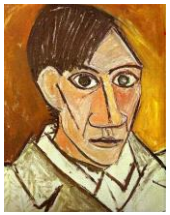


[Sethi, et al., ACM MM 2013]

Overview

- Reproducibility Crisis
- Addressing reproducibility with scientific workflows
- Case Study: Video Activity Recognition
- Case Study: Multimedia Analysis
- Case Study: Neural Algorithm of Artistic Style
- Benefits of scientific workflows for computer vision analysis
- Conclusions

Neural Algorithm of Artistic Style

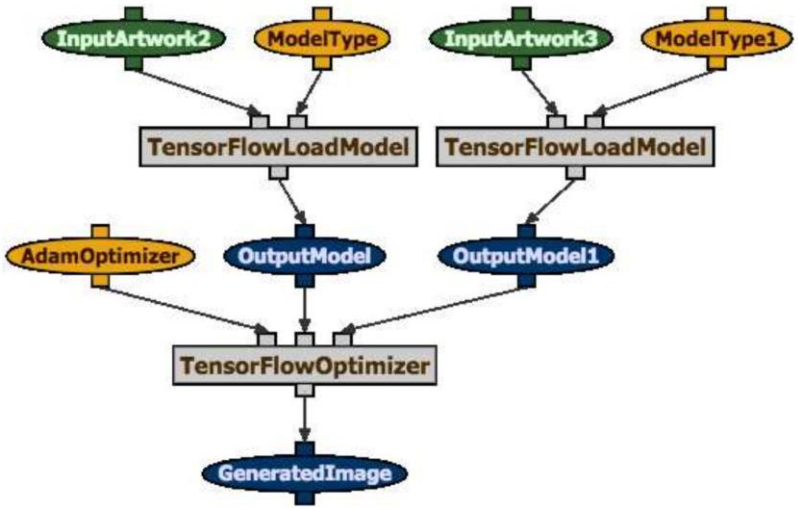
- The Neural Algorithm of Artistic Style by Gatys, et al., uses deep neural networks to separate the style and content of an image
  - Specifically, a Convolutional Neural Network, CNN
- Uses 2 images:
  - one image is a **style** image  and one is a *target* image 
- It then extracts the style from the **style** image and applies it to the content of the *target* image to create a new image in the style of the **style** image

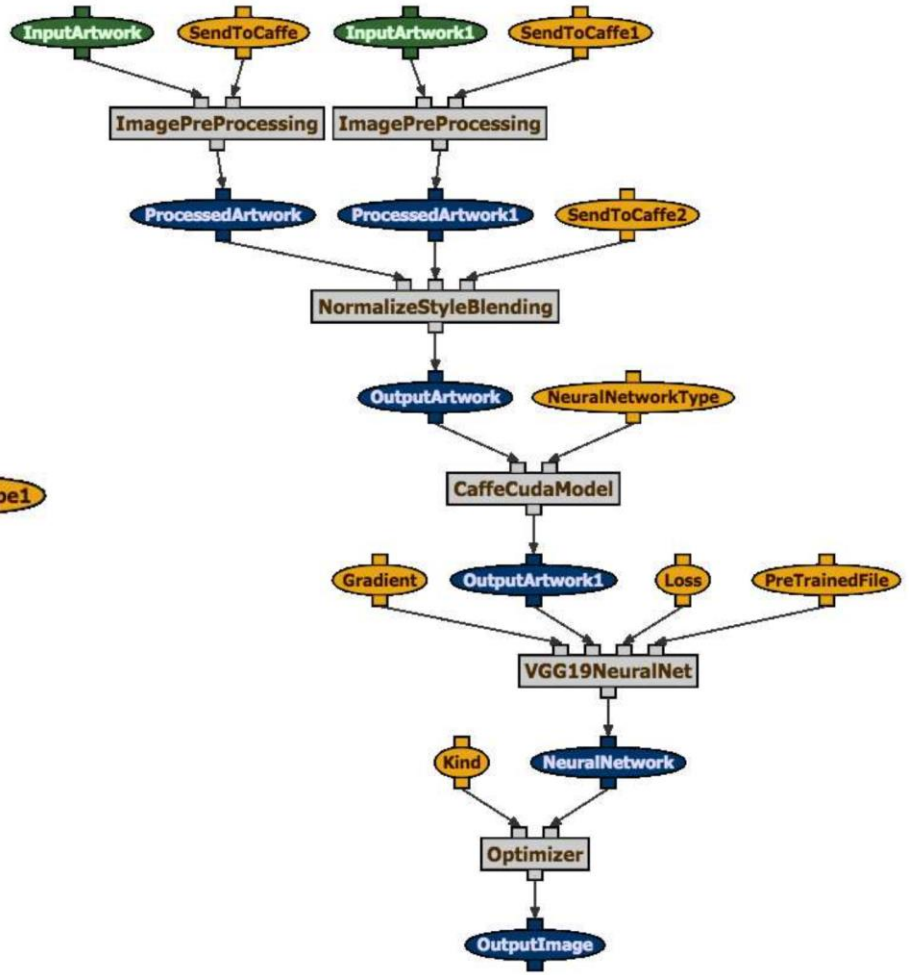
Reproducing their results

- We implemented two workflow versions: one using lua/torch and one using TensorFlow
- We reproduced the results from the paper
- We used the target image of a scene from Tubingen as presented in the original paper and reproduced their results as sho

Workflows

- Workflow using an implementation of CNNs that use the Lua/Torch languages
- Workflow using an implementation of CNNs that uses Google's TensorFlow library





Overview

- Reproducibility Crisis
- Addressing reproducibility with scientific workflows
- Case Study: Video Activity Recognition
- Case Study: Multimedia Analysis
- Case Study: Neural Algorithm of Artistic Style
- Benefits of scientific workflows for computer vision analysis
- Conclusions

Benefits of Workflows for computer vision analysis

- Accessibility
- Time savings
  - Site crawlers had been previously written, turned into workflow components in 2 days
  - Pre-existing workflows for text and video analytics: 1 day of work
  - Time/effort savings estimated at 300 hours of work
- Facilitate exploration and reuse
  - Explore different parameter values
  - Easy to add new components
  - Can use off-the-shelf components or roll your own

Conclusions

- Reproducibility in computer vision is challenging
- Collection of workflows and workflow fragments for computer vision
  - Quick deployment of state of the art techniques for image analysis
  - Integration of heterogeneous codebases and standard implementations
  - Easy to extend
- Future work: let non-experts to use image analysis workflows
  - Geoscience analysis of samples
  - Art students to analyze pieces of art