TRANSFORMING TRAIL CAMERA IMAGES OF STREAMS INTO A KNOWLEDGE BASE FOR ASSESSMENT

Mary Becker¹, Timothy Becker², Christopher Bellucci¹, Melissa Czarnowski³, Corinne Fitting³, Alexandria Hibbard³, Nicholas Muise² and Joseph Spada²

1. Monitoring and Assessment Program, CT Dept. of Energy and Environmental Protection, Hartford CT
2. Department of Computing Sciences, University of Hartford, West Hartford CT
3. Water Quantity Program, CT Dept. of Energy and Environmental Protection, Hartford CT

National CWA 303(d) and Data Management Training Workshop - June 10, 2021
STREAM FLOW MANAGEMENT

1979 Minimum Stream Flow Regulation, 1982 Diversion Act

2012 Adopted Stream Flow Regulations

Integrated Water Quality Reporting of Flow Alteration Impairments

State Water Plan Adopted in 2019
STREAM CONNECTIVITY

STREAM CONNECTIVITY CATEGORIES

Disconnected

Connected
STREAM CONNECTIVITY METRICS

DURATION
A period of time an image is associated with a category
Average number of consecutive days in category 1

FREQUENCY
How often an image is in a category
Number of days in category 1

MAGNITUDE
Provides a statistical summary of a category
Average flow category

TIMING
Describes when a category occurs temporally
Julian Day of 1st observation in category 1

Bar charts for different streams:
- (1) Bunnell Brook
- (2) Chidsey Brook
- (3) Cobble Brook
- (4) Honeypot Brook
- (5) Honeypot Brook
- (6) Mill River
- (7) Womenshenuck Brook

Legend:
- Blue: Connected
- Light Purple: Disconnected
- Red: Dry
BETTER ACCOUNTING OF FLOW ALTERATION

Flow Alteration Impairments 2010
34 Stream Miles
BETTER ACCOUNTING OF FLOW ALTERATION

Flow Alteration Impairments 2020
159 Stream Miles
MONITORING STREAM CONNECTIVITY WITH TRAIL CAMERAS

A few hundred images

• 2016 Trail Camera Sites
MONITORING STREAM CONNECTIVITY WITH TRAIL CAMERAS

A few thousand images

2017 Trail Camera Sites
MONITORING STREAM CONNECTIVITY WITH TRAIL CAMERAS

Hundreds of thousands of images

- 2021 Trail Camera Sites
- Regional Monitoring Network Sites
‘Conducting research is a bit like parenting. Raising a child involves a lot of cleaning and tidying, setting standards, and maintaining order, all of which goes completely unnoticed and for which the parent receives absolutely no credit.

Similarly, producing a bright, shiny result from the raw beginnings of a research project involves a lot of work that is almost never seen or acknowledged.’

-Paul Murrell ‘Introduction to Data Technologies’ (Emphasis Added)
DATA MANAGEMENT

Select Locations
- Capture at Least 1 Rifle-Pool Sequence
- Trim Vegetation

Inform Stream Management

Deploy Trail Camera
- Collect Hourly Images
  - Field Check and Download Images
  - Quality Control Image Data

Store and Organize Image Data
- Calibrate Image Readers

Calculate Average Daily Stream Connectivity Metrics
- Quality Control Image Stream Connectivity Category
- Store Category in Database

Assign Stream Connectivity Category to Images

Visualize Metrics
DATA WORKFLOW

- Metadata Organization & Cleaning
- Image Processing
- Data Import and QC
- Data Storage
- Stream Connectivity Labeling
STANDARDIZING METADATA

User Entered Folder Name

Folder Naming Convention
SiteID_SiteName_DeployStartDate_DeployEndDate_CameraID

Metadata Cleaning Tools

Python

Rename all image files

Flag image files that are taken outside of the deploy start/end dates

Check deploy date for daylight savings time and update times in exif data where needed

File Naming Convention
CAPTURING AND ORGANIZING DATA

SITE DETAILS

CAMERA INVENTORY

= Dependencies

DEPLOYMENT DETAILS

IMAGES

USER

CATEGORY DESCRIPTION

FLOW LABELS

MySQL
DATA IMPORT AND QUALITY CHECKS

Exchangeable Image File Format (EXIF)

```
'Image Copyright': (0x8298) ASCII=Copyright 2012 @ 286,
'Image DateTime': (0x0132) ASCII=2018:09:19 10:22:40 @ 266,
'Image ExifOffset': (0x8769) Long=830 @ 150,
'Image GPSInfo': (0x8825) Long=1668 @ 162,
'Image ImageDescription': (0x010E) ASCII=Moultrie Digital Game Camera @ 182,
'Image Make': (0x010F) ASCII=Moultrie @ 214,
'Image Model': (0x0110) ASCII=M-999i @ 230,
'Image Orientation': (0x0112) Short=35 @ 54,
'Image PrintIM': (0xC4A5) Undefined=[80, 114, 105, 110, 116, 73, 77, 45, 48, 51,
'Image ResolutionUnit': (0x0128) Short=Pixels/Inch @ 90,
'Image Software': (0x0131) ASCII=Ver 2.0 @ 258,
'Image XResolution': (0x011A) Ratio=96 @ 242,
'Image YCbCrPositioning': (0x0213) Short=Cohsited @ 126,
'Image YResolution': (0x011B) Ratio=96 @ 250,
```

Python

Automated Data Migration and Quality Check Tool

Extracts EXIF data (date/time, camera id, camera model)

Checks for correct naming structure, check file type and extracts data from file name

Checks for primary key and constraint violations

Image not meeting requirements are moved into a 'quarantine' folder and failure is logged in the database
IMAGE PROCESSING
MANUAL CLASSIFICATION: LABELING IMAGES
IMAGE CLASSIFICATION USING MACHINE LEARNING: DEEP LEARNING

large set of labeled images

m-class classification model
IMAGE CLASSIFICATION USING MACHINE LEARNING: DEEP LEARNING

new unlabeled image

m-class classification model

0: 0.0
1: 0.0
2: 0.2
3: 0.1
4: 0.1
5: 0.6
6: 0.0
If you randomly partition the test and training sets, you achieve a super high performance (that is also a terrible predictor given a new site)

(1) Starting with the lowest frequency category, select sites such that they keep the highest consistency with respect to the test and training partition spectrums (select all of site 15244 images)

(2) Select all of a category for a site (so that the training never sees the category 1 from site 15244)
M-CLASS TRAINING: BALANCING

When label spectrums are biased, the training will not learn less frequent categories effectively.

(1) Down sample higher frequency labels

(2) Up sample lower frequency labels (requires image manipulation code to on-the-fly alter the images which is known as data augmentation)

sid=15244

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:244</td>
<td></td>
</tr>
<tr>
<td>2:203</td>
<td></td>
</tr>
<tr>
<td>3:297</td>
<td></td>
</tr>
<tr>
<td>4:1669</td>
<td></td>
</tr>
<tr>
<td>5:147</td>
<td></td>
</tr>
<tr>
<td>6:6</td>
<td></td>
</tr>
</tbody>
</table>
CLASS TRAINING: 2-CLASS F1 = 0.73
Thanks to...

**CT DEEP field assistants and program staff**

**U.S. EPA** for providing equipment that aids the collection of data for this study

**University of Hartford** for providing grant funding for GPU resources needed for image processing and model development