Title

Road Damage Detection and Classification with Detectron2 and Faster R-CNN

Authors

Vung Pham, Chau Pham, Tommy Dang

Publication date

2020/12/10

Conference

2020 IEEE International Conference on Big Data (Big Data)

Pages

5592-5601

Description

The road is vital for many aspects of life, and road maintenance is crucial for human safety. One of the critical tasks to allow timely repair of road damages is to quickly and efficiently detect and classify them. This work details the strategies and experiments evaluated for these tasks. Specifically, we evaluate Detectron2's implementation of Faster R-CNN using different base models and configurations. We also experiment with these approaches using the Global Road Damage Detection Challenge 2020, A Track in the IEEE Big Data 2020 Big Data Cup Challenge dataset. The results show that the X101-FPN base model for Faster R-CNN with Detectron2's default configurations is efficient and general enough to be transferable to different countries in this challenge. This approach results in F1 scores of 51.0% and 51.4% for the test1 and test2 sets of the challenge, respectively. Though the visualizations show good prediction results, the F1 scores are low. Therefore, we also evaluate the prediction results against the existing annotations and discover some discrepancies. Thus, we also suggest strategies to improve the labeling process for this dataset.

Total citations

Cited by 169

Title

Road Damage Detection and Classification with YOLOv7

Authors

Vung Pham, Du Nguyen, Christopher Donan

Publication date

2022/12

Conference

2022 IEEE International Conference on Big Data (Big Data)

Description

Maintaining the roadway infrastructure is one of the essential factors in enabling a safe, economic, and sustainable transportation system. Manual roadway damage data collection is laborious and unsafe for humans to perform. This area is poised to benefit from the rapid advance and diffusion of artificial intelligence technologies. Specifically, deep learning advancements enable the detection of road damages automatically from the collected road images. This work proposes to collect and label road damage data using Google Street View and use YOLOv7 (You Only Look Once version 7) together with coordinate attention and related accuracy fine-tuning techniques such as label smoothing and ensemble method to train deep learning models for automatic road damage detection and classification. The proposed approaches are applied to the Crowdsensing-based Road Damage Detection Challenge (CRDDC2022), IEEE BigData 2022. The results show that the data collection from Google Street View is efficient, and the proposed deep learning approach results in F1 scores of 81.7% on the road damage data collected from the United States using Google Street View and 74.1% on all test images of this dataset. With these results, we received rank 2 (silver prize) as a data contributor and rank 3 (bronze prize) as the predictive model in this competition among 54 leaders (private companies and academic institutions) in this area.

Total citations

Cited by 55

Title

Cvexplorer: Multidimensional visualization for common vulnerabilities and exposures

Authors

Vung Pham, Tommy Dang

Publication date

2018/12/10

Conference

2018 IEEE International Conference on Big Data (Big Data)

Pages

1296-1301

Publisher

IEEE

Description

Cyber attacks cause great damage to our national security, ranging from individual internet user to biggest governmental/industrial organizations, such as Equifax (Data Breach 145.5 Million Accounts, reported in July 2017) or Uber (Data Breach 57 Million Records, reported in November 2017). The cyber assault has significantly increased in breadth and depth. This paper introduces CVExplorer, a novel interactive system for visualizing cybersecurity threats reported in the National Vulnerability Database. The proposed system aims to work as a reporting and alerting tool that can help enhance the security against cyber attacks can potentially reduce network vulnerabilities. The CVExplorer system containing multiple linked views allows users to visualize the relationships of various dimensions in the large number of vulnerability reports, such as types and levels of vulnerability, vendors, and products. The CVExplorer provides an intuitive interface and supports a range of interactive features, such as filtering and ordering by vulnerability severity ratings, allowing users to narrow down topics of interest quickly. To demonstrate the effectiveness of the proposed system, we demonstrate the CVExplorer on two case studies of Common Vulnerabilities and Exposures retrieved from the National Vulnerability Database.

Total citations

Cited by 51

Title

Enhanced soil profile visualization using portable X-ray fluorescence (PXRF) spectrometry

Authors

Fujun Sun, Noura Bakr, Tommy Dang, Vung Pham, David C Weindorf, Zhuodong Jiang, Hualei Li, Qiu-Bing Wang

Publication date

2020/1/15

Journal

Geoderma

Volume

358

Pages

113997

Publisher

Elsevier

Description

Soil horizonation has traditionally relied upon morphological description, field sampling, and laboratory analysis as a means of establishing key diagnostic horizons and features within a soil profile. While this technique is useful for many soil properties, other quantifiable properties associated with elemental abundance may be visually imperceptible. Portable X-ray fluorescence (PXRF) spectrometry allows for rapid elemental data acquisition in-situ, providing more than 20 elements commonly detectable within a soil profile. In this study, a grid containing 130 cells was physically imposed upon two soil profiles in West Texas, USA. PXRF was used to sequentially scan each grid cell and acquire elemental data. The profiles were morphologically described by a panel of soil scientists, sampled, and subjected to traditional laboratory characterization. Data from the PXRF and field sampling were statistically compared, then visualized as depth functions. A data visualization tool was created to overly elemental data on top of a digital photograph of the soil profile in semitransparency. Using elemental data as a proxy, PXRF data was able to detect both argillic and calcic horizons within one soil profile and visually illustrate their extent and assist in delineating horizons in a second profile with nondescript horizons. The soil visualization tool has other features as well such as correlation analysis between elements, adjustable correlation threshold in real time, rasterized vs. smoothed data overlay, and box-plots of elemental concentrations. Summarily, this research developed a new tool for PXRF data visualization in support of soil pedon description, allowing for elemental data which may be visually imperceptible to be considered during soil profile characterization. Future advancements will seek to integrate this tool directly into the PXRF software for on-site visualization of any element specified by the analyst. Applications of such could include real-time mapping of contaminant metals in soil profiles.

Total citations

Cited by 43

Title

Soil profile analysis using interactive visualizations, machine learning, and deep learning

Authors

Vung Pham, David C Weindorf, Tommy Dang

Publication date

2021/12/1

Journal

Computers and Electronics in Agriculture

Volume

191

Pages

106539

Publisher

Elsevier

Description

Soil is an essential element of life, and soil properties are crucial in analyzing soil health. Recent developments of proximal sensor technologies, such as portable X-ray fluorescence (pXRF) spectroscopy or visible and near-infrared (Vis–NIR) spectroscopy, offer rapid and non-destructive alternatives for quantifying data from soil profiles. While the data collection time using these technologies decreases significantly, the subsequent analysis remains time-consuming, and current analysis solutions only provide basic visualizations. Furthermore, the use of collected data from proximal sensors to predict high-level soil properties has garnered worldwide attention in the past decade, owing to its convenience. Therefore, this paper discusses the objectives for software solutions in this area, consolidated from interviewing 102 stakeholders. Following these requirements, data visualizers work closely with soil scientists to propose a set of interactive visualizations for analyzing soil profiles using pXRF data. These interactive visualizations receive positive feedback from the domain experts. This project also explores various machine learning and deep learning approaches to predict soil properties from spectral data. This work then proposes a deep learning model called RDNet that achieves state-of-the-art results in predicting pHH2O and pHKCL from Vis–NIR spectra acquired from a set of globally distributed soil samples.

Title

WordStream: Interactive Visualization for Topic Evolution.

Authors

Tommy Dang, Huyen N Nguyen, Vung Pham, J Johansson, F Sadlo, GE Marai

Publication date

2019

Conference

EuroVis (Short Papers)

Pages

103-107

Description

This paper introduces WordStream, an interactive visual tool for the demonstration of topic evolution. Our approach utilizes the two popular techniques. Word clouds are designed to give an engaging visualization of text via font sizes and colors, while stacked graphs are a common method for visualizing topic evolution. In particular, WordStream emphasizes essential terms chronologically and spatially. To show the usefulness of WordStream, we demonstrate its applications on various data sets, including the Huffington Post and IEEE VIS publications.

Total citations

Cited by 29

Title

Mtsad: Multivariate time series abnormality detection and visualization

Authors

Vung Pham, Ngan Nguyen, Jie Li, Jon Hass, Yong Chen, Tommy Dang

Publication date

2019/12/9

Conference

2019 IEEE international conference on big data (Big Data)

Pages

3267-3276

Publisher

IEEE

Description

Detecting outliers is one of the fundamental tasks in visual analytics and valuable in many application domains, such as suspicious network cyberattack recognition. This paper introduces an approach to analyzing and visualizing high-dimensional time series, focusing on identifying multivariate observations that are significantly different from the others. We also propose a prototype, called MTSAD, to guide users when interactively exploring abnormalities in large time series. The prototype contains two views: the main window provides an overview of identified outliers overtime, the detail window investigates and explores the ranked temporal data entries based on their outlying contributions to the overall plots. The visual interface supports a full range of interactions, such as lensing, brushing and linking, ranking, and filtering. To validate the benefits and usefulness of our approach, we demonstrate MTSAD on real-world datasets of different numbers of attributes.

Title

Deepvix: Explaining long short-term memory network with high dimensional time series data

Authors

Tommy Dang, Hao Van, Huyen Nguyen, Vung Pham, Rattikorn Hewett

Publication date

2020/7/1

Book

Proceedings of the 11th international conference on advances in information technology

Pages

1-10

Description

Machine learning automates the process of analytical model building by means of the computing power of machines. Visual analytics couples interactive visual representations and underlying analysis, putting the human at the center of the analytics and decisionmaking process. This paper aims to combine the strengths of both data science fields into a unified system, called DeepVix, which focuses on the visual explainability of the multivariate time-series predictions using neural networks. Within our DeepVix system, a visual presentation of the neural network explains the intermediate steps, as well as the temporal weights of various gates of the entire learning process. The relationships between input variables and the target variable can also be inferred automatically from the trained model. Interactive operations allow users to explore the neural network, to gain understandings of the model and essential features with layers and nodes, and finally to customize the neural network configurations to fit their needs. We demonstrate our approach with Recurrent Deep Learning on various real-world time series datasets, including the multivariate measurements of a medium-size High-Performance Computing Center, the S&P500 stock data over the past 39 years, and the US employment data retrieved from the Bureau of Labor and Statistics.

Total citations

Cited by 25

Title

Visualization and explainable machine learning for efficient manufacturing and system operations

Authors

Dy D Le, Vung Pham, Huyen N Nguyen, Tommy Dang

Publication date

2019/12/1

Journal

Smart and Sustainable Manufacturing Systems

Volume

3

Issue

2

Pages

127-147

Publisher

ASTM International

Description

To enable Industry 4.0 successfully, there is a need to build a resilient automation system that can quickly recover after having been attacked or robustly sustain continued operations while being threatened, enable an automated monitoring evolution via various sensor channels in real time, and use advanced machine learning and data analytics to formulate strategies to mitigate and eliminate faults, threats, and malicious attacks. It is envisioned that if we can develop an intelligent model that (a) represents a meaningful, realistic environment and complex entity containing manufacturing Internet of Things interdependent and independent properties that are stepping-stones of the cyber kill chain or precursors of the onset of cyberattacks; (b) can learn and predict potential errors and formulate offense/defense strategies and healing solutions; (c) can enable cognitive ability and human-in-the-loop analytics in real time; and (d) can facilitate system behavior changes to disrupt the attack cascade, then the hosting system can learn how to neutralize threats and attacks and self-repair infected or damaged links autonomously. In this article, our preliminary work presents a visual analytics framework and technique for situational awareness, including autonomously monitoring, diagnosing, and prognosticating the state of cyber-physical systems. Our approach, presented in this article, relies on visual characterizations of multivariate time series and real-time predictive analytics to highlight potential faults, threats, and malicious attacks. To validate the usefulness of our approach, we demonstrate the developed technique using various aviation datasets obtained from the Prognostics Center of Excellence at the National Aeronautics and Space Administration Ames.

Total citations

Cited by 24

Title

RDD2022-The multi-national Road Damage Dataset released through CRDDC'2022

Authors

D Arya, Hiroya Maeda, Yoshihide Sekimoto, Hiroshi Omata, Sanjay Kumar Ghosh, Durga Toshniwal, Madhavendra Sharma, Van Vung Pham, Jingtao Zhong, Muneer Al-Hammadi, Mamoona Birkhez Shami, Du Nguyen, Hanglin Cheng, Jing Zhang, Alex Klein-Paste, Helge Mork, Frank Lindseth, Toshikazu Seto, Alexander Mraz, Takehiro Kashiyama

Publication date

2022/1/1

Journal

(No Title)

Publisher

figshare

Description

Description The Road Damage Dataset, RDD2022, is released as a part of the Crowdsensing-based Road Damage Detection Challenge (CRDDC'2022), an IEEE BigData Cup. It comprises 47,420 road images from six countries, Japan, India, the Czech Republic, Norway, the United States, and China. The images have been annotated with more than 55,000 instances of road damage. Four types of road damage, namely longitudinal cracks, transverse cracks, alligator cracks, and potholes, are captured in the dataset. Usage The annotated dataset is envisioned for developing deep learning-based methods to detect and classify road damage automatically. The municipalities and road agencies may utilize the RDD2022 dataset, and the models trained using RDD2022 for low-cost automatic monitoring of road conditions. Further, computer vision and machine learning researchers may use the dataset to benchmark the performance of different algorithms for other image-based applications of the same type (classification, object detection, etc.). For further details, please refer to the CRDDC'2022 resources.

Total citations

Cited by 17

Title

Scagcnn: Estimating visual characterizations of 2d scatterplots via convolution neural network

Authors

Vung Pham, Ngan VT Nguyen, Tommy Dang

Publication date

2020/7/1

Book

Proceedings of the 11th international conference on advances in information technology

Pages

1-9

Description

Scagnostics is a set of visual features that characterizes the data distribution of a 2D scatterplot and has been used in a wide range of applications. However, calculating the scagnostics scores involves computationally expensive algorithms. Moreover, the algorithms are sensitive to the slight changes in the underlying data distribution within the scatterplot. Therefore, this work provides a machine learning model, called ScagCNN, to estimate the scagnostics scores. This model aims to improve the scagnostics computation time and to reduce the sensitivity to the small shifts in the data distribution. This work also provides a web prototype to explore the predictive performance of the model and to give a visual explanation about whether a prediction is accurate. Furthermore, we test the performance of our solution on datasets of various sizes.

Total citations

Cited by 11

Title

Finanviz: Visualizing emerging topics in financial news

Authors

Ngan VT Nguyen, Vinh T Nguyen, Vung Pham, Tommy Dang

Publication date

2018/12/10

Source

2018 IEEE International Conference on Big Data (Big Data)

Pages

4698-4704

Publisher

IEEE

Description

The explosion of social media has paved a way for big data in which entrepreneurs use this data to find out potential customers, market demands, individual behavior, thereby to improve existing products, to create new products according to users' need, or to analyze and evaluate financial risks. The challenges of the heterogeneity and fragmentation of data make it difficult for analysts to fully exploit the benefit of deluge information. Available statistical software lacks customization and address unknown research questions. This paper proposes FinanViz, a visual analytics tool for analyzing financial news on social media. The principal aim of FinanViz is to observe the dynamic behavior of terms/words over time along with their proximity to other terms/words. The tool provides an intuitive, interactive exploration of the financial topics and what events are emerging in which we would argue that it will give hints for financial marketers in the decision making process.

Total citations

Cited by 11

Title

SOAViz: Visualization for portable X-ray fluorescence soil profiles

Authors

Vung Pham, Tommy Dang

Publication date

2019

Publisher

The Eurographics Association

Description

The soil is an essential element of life. It is where people grow plants for food, fibers, and other materials. It also helps to filter water and recycles wastes. Therefore, understanding soil physical/chemical characteristics and structural aggregation are of vital importance. In this project, we work closely with the soil scientists to develop a visualization solution to the rapidly gaining favor approach to soil horizon analysis using Portable X-ray Fluorescence (pXRF) devices. Our visualization, called SOAViz, aims to provide soil scientists with rapid valuable insights into soil properties both visually perceptible with graphs and imperceptible quantification features with statistical calculations from the data collected from pXRF equipment. SOAViz was developed with analysis tasks solicited from the soil scientists and

validated by applying to real soil profiles collected in an Experimental Rangeland in Lubbock, TX, USA. This visual solution together with the quick scanning results from pXRF devices offers a timely means of quantifying elemental concentrations in the soil horizons in large scale at a reduced cost.

Total citations

Cited by 10

Title

Mtdes: Multi-dimensional temporal data exploration system

Authors

Vung V Pham, Tommy Dang

Publication date

2018/10/21

Conference

2018 IEEE Conference on Visual Analytics Science and Technology (VAST)

Pages

100-101

Publisher

IEEE

Description

This work proposes a visual analytic solution which is well-designed to provide investigative functions with fluent interactions to analyze multi-dimensional temporal data. The solution allows users to view different dimensions of the data at different levels of details with a well-designed mixture of different visualizations and smooth interactions. At the general/overview level, various aggregation strategies are used to reduce data to be visualized, and different sorting procedures are used to cluster correlated data together to help discover patterns. Detail views are provided to explore and confirm/reject the identified patterns. Interaction and smooth transition between views are implemented to enable natural actions while performing analysis tasks. This work also presents the result of applying the solution to the VAST 2018 - Mini-Challenge (MC) 2 dataset, which led to the Strong Support for Exploratory Analysis award for the challenge.

Total citations

Cited by 10

Title

AgasedViz: visualizing groundwater availability of Ogallala Aquifer, USA

Authors

Tommy Dang, Vung Pham, Huyen N Nguyen, Ngan VT Nguyen

Publication date

2020/3

Journal

Environmental Earth Sciences

Volume

79

Issue

5

Pages

110

Publisher

Springer Berlin Heidelberg

Description

Water plays an immensely important role in human life, from household to agriculture and industrial development. Due to the heavy utilization and growing demands of water, the availability of water has become a high priority. Besides sources of surface water, such as streams, lakes, and rivers, aquifers are a significant source of water on Earth through the essential groundwater supply. In the United States, the Ogallala Aquifer, a primary geologic unit of the High Plains Aquifer System, is under massive exploitation, and the water level has decreased substantially. Analysis of the saturated thickness of this aquifer is urgently essential. Thus, an interactive visual analytics tool is necessary to enable users to visualize water availability and variation overtime at multiple locations within the region. Our

interactive analytic tool carries out this by first retrieving and processing the data derived from the saturated thickness of groundwater using sensors integrated into the wells. Subsequently, a visualization consisting of a contour map and a time series heatmap is created based on the sensor data, to capture the trends and patterns, such as underground water distribution, spatial and temporal changes, and sudden decreases or increases of the water level. The visual components allow users to explore observational data, which organize the visual structure in supporting making an inference to gain insights. insights. This approach can be extended to apply for any geographic areas for water-level monitoring and controlling.

Total citations

Cited by 9

Title

Visualization of data from HACC simulations by Paraview

Authors

Bao D Nguyen, Ngan VT Nguyen, Vung Pham, Tommy Dang

Publication date

2019/10/20

Conference

2019 IEEE Scientific Visualization Conference (SciVis)

Pages

31-32

Publisher

IEEE

Description

In this project, we visualize and analyze data of Hybrid Accelerated Cosmology Code (HACC) cosmology simulations. We investigate in point rendering as well as Smoothed-particle hydrodynamics (SPH) interpolation and apply them to visualize the data in a 3D environment. Paraview was used for the rendering task. Some parameters of points in the software were adjusted to find well-qualified visualizations. These approaches provide a clear vision to the impact of Active Galactic Nuclei (AGN) on surrounding matter, and thus, we can gain more information for understanding the nature of AGNs and the universe.

Total citations

Cited by 7

Title

MalView: Interactive visual analytics for comprehending malware behavior

Authors

Huyen N Nguyen, Faranak Abri, Vung Pham, Moitrayee Chatterjee, Akbar Siami Namin, Tommy Dang

Publication date

2022/9/19

Journal

IEEE Access

Volume

10

Pages

99909-99930

Publisher

IEEE

Description

Malicious applications are usually comprehended through two major techniques, namely static and dynamic analyses. Through static analysis, a given malicious program is parsed, and some representative artifacts (e.g., control-flow graphs) are produced without any execution; whereas, the given malicious application needs to be executed when conducting dynamic analysis. These two mainstream techniques for analyzing the given software are effective in detecting certain classes of malware. More specifically, through static analysis, the patterns and signature of the malware are exposed, helping in detecting any known malicious payload hidden in or injected into the code. On the other hand, behavioral and run-time execution patterns of software are explored through dynamic analysis. To ease the analysis process, a third analysis approach, known as the visual representation of the artifacts created by both static and dynamic analysis tools, would also be a supplementary asset for malware experts. This paper introduces MalView, an interactive visualization

platform, for malware analysis by which pattern matching techniques on both signature-based and behavioral analysis artifacts can be utilized to 1) classify malware, 2) identify the intention and location of the malicious payload in the artifacts, 3) analyze unknown malware (i.e., zero-day malware) by recognizing any unusual signature or behavior, and 4) explore the time dependencies and thus the system components affected or tampered by the underlying malware. The results of several case studies conducted in this work show that MalView offers more features and information compared to some other visualization tools, facilitating the malware analysis process.

Total citations

Cited by 6

Title

Graph adversarial attacks and defense: An empirical study on citation graph

Authors

Chau Pham, Vung Pham, Tommy Dang

Publication date

2020/12/10

Conference

2020 IEEE International Conference on Big Data (Big Data)

Pages

2553-2562

Publisher

IEEE

Description

This paper details the methodologies and decisions making processes used while developing the attacking and defending models for the Graph Adversarial Attacks and Defense applied to a large citation graph. To handle the large graphs, our attack strategy is twofold: 1) randomly attack the structure first, 2) keep the structure unchanged, then continue the attack on the features using the gradient-based method. On the other hand, the defender is based on 1) filtering and normalizing the feature data, 2) applying the Graph Convolutional Network model, and 3) selecting the models with the highest accuracy and robustness based on our

own attacking data. We applied these strategies in KDD Cup 2020 on Graph Adversarial Attacks and Defense dataset. The attacker can drop the accuracy of a surrogate 2-layer Graph Convolutional Network model from 60% to 30% on the test set. Our defending model has 68% accuracy on the validated data and has 89% of the target labels remained the same while adding fake nodes, generated by our attacking method, to the graph.

Total citations

Cited by 6

Title

ScagnosticsJS: Extended Scatterplot Visual Features for the Web

Authors

Vung Pham, Tommy Dang

Publication date

2020

Conference

EUROGRAPHICS 2020

Description

Scagnostics is a set of features that characterizes the data distribution in a scatterplot. These visual features have been used in various applications to detect unusual correlations of bivariate data. However, there is no formally published implementation for 3D or higher. This project aims to provide the Scagnostics implementation in JavaScript, called ScagnosticsJS, and also extend these measures for higher dimensional scattered points. We also present a Scagnostics exploration webpage, which makes the underlying algorithms transparent to users.

Total citations

Cited by 6

Title

Solar flare prediction using two-tier ensemble with deep learning and gradient boosting machine

Authors

Chau Pham, Vung Pham, Tommy Dang

Publication date

2019/12/9

Conference

2019 IEEE International Conference on Big Data (Big Data)

Pages

5844-5853

Publisher

IEEE

Description

This paper describes a machine learning approach to the solar flare prediction competition, a track in IEEE Big Data 2019 Big Data Cup. The competition task is to predict whether or not there is a solar flare event basing on a given time series of solar magnetic field parameters. Our method involves exploring and constructing data-driven machine learning models for the classification task of two imbalanced class labels from time series. Specifically, the investigated models include boosting, logistic regression, multilayer perceptron neural network, and long short-term memory neural network. These models have been successfully deployed and combined in an ensemble framework with two tiers in our final proposed solution for this competition. Our proposed approach ranked at the second place in the competition (the first on the private board and the eleventh on the public board).

Total citations

Cited by 6

Title

HealthTvizer: Exploring health awareness in twitter data through coordinated multiple views

Authors

Tommy Dang, Ngan VT Nguyen, Vung Pham

Publication date

2018/12/10

Conference

2018 IEEE International Conference on Big Data (Big Data)

Pages

3647-3655

Publisher

IEEE

Description

Analyzing public user posts and shared information on social media can assist us in measuring various population characteristics, patterns, movements, and as well as the public health conditions. In recent years, researchers have been analyzing social media (such as Facebook or Twitter feeds) to detect and predict various emerging events and market trends. Fewer attentions have been paid to the epidemic of the diseases. In this paper, we present a social media analytics tool, called HealthTvizer, for exploring health awareness using Twitter data through interactive and interconnected multiple views. We use topic modeling to pick the relevant and meaningful terms from more than 57 million tweets. We detect the disease name and related contents which are shared by the users of different geographical locations (mostly in the United States). We believe that the collected geolocations from the users' tweets can reveal the patterns of diseases for a given term which allows a researcher to detect, analyze, and explore information about the diseases and hence take necessary steps to improve public health awareness. We validate the effectiveness of HealthTvizer through an informal user study. The feedback from this study also motivates us on interesting future extensions of the tool.

Total citations

Cited by 6

Title

IoTViz: Visualizing emerging topics in the internet of things

Authors

Vung Pham, Vinh T Nguyen, Tommy Dang

Publication date

2018/12/10

Source

2018 IEEE International Conference on Big Data (Big Data)

Pages

4569-4576

Publisher

IEEE

Description

The "Internet of Things" is changing the way companies operate and consumers behave. Therefore, it is essential to capture trends in "Internet of Things". This paper proposes IoTViz, a visual analytics tool for analyzing "Internet of Things" news on social media. The principal aim of IoTViz is to observe the dynamic behavior of topics along with their proximity to other dimensions such as user comments and ratings in multiple coordinated views. IoTViz provides an interactive exploration of the IoT topics and supports of a range of interactive features, such as linking and filtering, allowing users to narrow down events of interest quickly. It is interesting to filter and visualize IoT news regarding the individual organization, e.g., user opinions/ratings regarding a company or its products.

Total citations

Cited by 6

Title

Outliagnostics: Visualizing Temporal Discrepancy in Outlying Signatures of Data Entries

Authors

Vung Pham, Tommy Dang

Publication date

2019

Conference

2019 IEEE Visualization in Data Science (VDS)

Pages

29-37

Description

This paper presents an approach to analyzing two-dimensional temporal datasets focusing on identifying observations that are significant in calculating the outliers of a scatterplot. We also propose a prototype, called Outliagnostics, to guide users when interactively exploring abnormalities in large time series. Instead of focusing on detecting outliers at each time point, we monitor and display the discrepant temporal signatures of each data entry concerning the overall distributions. Our prototype is designed to handle these tasks in parallel to improve performance. To highlight the benefits and performance of our approach, we illustrate and validate the use of Outliagnostics on real-world datasets of various sizes in different parallelism configurations. This work also discusses how to extend these ideas to handle time series with a higher number of dimensions and provides a prototype for this type of datasets.

Total citations

Cited by 5

Title

SoilScanner: 3D Visualization for Soil Profiling using Portable X-ray Fluorescence

Authors

V Pham, D Weindorf, T Dang

Conference

The Eurographics Association

Description

Soil scientists perform similar types of exploratory analysis repeatedly, such as generating the spatial distribution of chemical elements. The soil analysis process is time-consuming (may take days or weeks), labor-intensive (involving many people with different expertise for data collection, measurements, visual representation, and data analysis), and involving various tools (from traditional software, such as Microsoft Excel, to some complicated packages such as ArcGIS and MatLab). Inspired by medical scanning, this paper proposes a 3D visual solution, which can be generated via a web interface, allowing Soil scientists to perform on-the-field analysis. Our visualization prototype, named SoilScanner, supports a full range of interactive operations, such as ranking, filtering, brushing and linking, and detail on demand. We also demonstrated the usability of our SoilScanner visualizations on the soil profiles in West Texas, USA, collected via portable X-ray fluorescence spectrometers.

Total citations

Cited by 5

Title

Contimap: Continuous heatmap for large time series data

Authors

Vung Pham, Ngan Nguyen, Tommy Dang

Publication date

2020/10/26

Conference

2020 Visualization in Data Science (VDS)

Pages

42-51

Publisher

IEEE

Description

Limited human cognitive load, limited computing resources, and finite display resolutions are the major obstacles for developing interactive visualization systems in large-scale data analysis. Recent technological innovation has significantly improved computing power, such as faster CPUs and GPUs, as well as display resources, including ultra-high-resolution displays and video walls. However, large and complex data is still ahead in the run as we are generating huge amounts of data daily. Our strategy to bridge these gaps is to present the right amount of information through the use of compelling graphics. This paper proposes an approximation algorithm and a web prototype for representing a high-level abstraction of time series based on heatmap designs. Our approach aims to handle a significant amount of time series data arising from various application domains, such as cybersecurity, sensor network, and gene expression analysis.

Total citations

Cited by 4

Title

Ufo_tracker: Visualizing ufo sightings

Authors

Vinh T Nguyen, Vung Pham, Tommy Dang

Publication date

2018/12/10

Conference

2018 IEEE International Conference on Big Data (Big Data)

Pages

4352-4359

Publisher

IEEE

Description

Visualizing and analyzing geospatial and temporal observations are common tasks for many application domains. In this paper, we introduce UFO Tracker, a visual analytic tool for analyzing unidentified flying object sightings from the National UFO Reporting Center. The goal here is to give the user a higher level view of where different types of sightings occur, to investigate whether sightings are increasing or decreasing over time, to discover the connections between different events which might happen at different geographic areas, and to quickly identify typical incidents at a given period of time without reading the whole sightings through topic modelling. Multiple visualization and data mining techniques are combined to make sense the increasingly large UFO reports which get updated hourly. The usefulness of the application is evaluated through a case study where anon-expert in ufology can find some typical interesting sightings. Our application can also be able to detect some misleading events such as missile launch or fireworks on a specific day through keywords and topic extraction. One limitation of our application is the data which is not up-todate when new sightings are posted since the application pulled and processed data locally. Our initial application targets UFO sighting reports. However, we believe our approach has wider applications in other research domains, such as analyzing text corpus obtained from social media.

Total citations

Cited by 4

Title

iDVS: interactive 2D and 3D visualizations of proximal sensor data for rapid characterization of soil profiles

Authors

Vung Pham, Cynthia M Jordan, Matthew G Siebecker, David C Weindorf, Tommy Dang

Publication date

2023/4

Journal

Precision Agriculture

Volume

24

Issue

2

Pages

627-646

Publisher

Springer US

Description

Knowledge of the soil’s physical and chemical properties in field-scale geographical areas is crucial for farmers and policymakers for agronomic productivity and environmental quality assessment. Proximal sensors can successfully model soil properties for these purposes and offer a way to rapidly acquire data from soil profiles. However, existing data analysis approaches are ill-suited to explore this type of multivariate proximal sensor data over large land areas and in a sizeable three-dimensional volume. Therefore, this work proposes a multifaceted approach with seamless integration of a star pattern for soil sample collection, data acquisition using proximal sensor devices, and an interactive data visualization solution for processing, analyzing, and reporting analysis results. This solution is the result of an interdis- ciplinary project in which data visualizers worked closely with soil scientists and agronomists to develop a tool called iDVS for rapid characterizations of soil profiles over larger geographical areas using proximal sensor technologies.

Total citations

Cited by 3

Title

DualNetView: Dual Views for Visualizing the Dynamics of Networks

Authors

Vung Pham, Ngan Nguyen, Tommy Dang

Publication date

2020/5

Conference

EuroVis Workshop on Visual Analytics (2020)

Publisher

Eurographics Proceedings 2020, The Eurographics Association.

Description

The force-directed layout is a popular visual method for revealing network structures, such as clusters and important vertices. However, it is not capable of representing temporal patterns, such as how clusters/communities evolve. Dynamic network visualizations trade the overall structures for temporal relationships. In this paper, we present a dual view framework for capturing both overall structures and temporal patterns within networks. The linked supplemental views utilize the strengths of both visualization techniques to provide useful insights into the given networks. To demonstrate the usefulness of our proposed dual views, we provide three use cases of dynamic networks: computer networks communications, activities of suspicious processes in computer systems, and social networks.

Total citations

Cited by 3

Title

IoTNegViz: An interactive tool for visualizing negative aspects of IoT

Authors

Huyen N Nguyen, Vinh T Nguyen, NV Nguyen, Vung Pham, Tommy Dang

Publication date

2018/12

Journal

2018 IEEE International Conference on Big Data (Big Data)

Pages

4565-4568

Description

The “Internet of Things”, or IoT, brings to people a brand new approach with a wide range of applications of how we live in this ever changing world. Besides various advantages of IoT that users can benefit from, the downsides of such cyber physical system are often neglected. This paper introduces IoTNegViz, an analytic tool for visualizing possible negative aspects from a cyber physical system in IoT. First, the data is extracted from Twitter, then a natural language processing tool is applied for generating and categorizing keywords. The visualization is utilized for exploring and capturing information in a spatial manner. To evaluate the usefulness and efficiency of IoTNegViz, we conduct a use-case to gather feedback and experience form users.

Total citations

Cited by 3

Title

Malware API call-based multiclass-classification using machine learning and deep learning

Authors

Sarah Adair, Cihan Varol, Fan Liang, Vung Pham

Publication date

2024/4/29

Conference

2024 12th International Symposium on Digital Forensics and Security (ISDFS)

Pages

1-7

Publisher

IEEE

Description

Malicious attacks have been on the rise with the growth of technological innovations. With the increase in mali-cious attacks, many current defense systems focus on preventive and reactive measures. However, the Intrusion prevention system and Intrusion Detection system may not be able to detect all suspicious processes and files that come through the network. Even if the security system detects suspicious activity, it may be unable to quickly classify the type of malware to implement effective and efficient protection. Thus, much research has focused on creating machine-learning models to classify suspicious programs as malicious or benign. However, the limitation of classifying a suspicious file as benign or malicious still does not aid the system in deploying effective security measures to counteract the damage of executed suspicious files. We proposed to train machine learning models and deep learning models to classify the eight most common malware classes. Trained models that can classify malware classes will boost a system's efficiency in identifying the type of mal ware, focusing on the affected area, and deploying effective countermeasures to minimize damages. Our experiment shows that the XGBoost model performs the best for tabular data types with an average accuracy of 65%, and the Transformer model achieves the highest accuracy score of 57 % for sequential datasets.

Total citations

Cited by 2

Title

Deep learning approaches for fingerprint verification

Authors

Nikita Dalvi, Van Vung Pham

Publication date

2023/10/15

Publisher

Institute for Homeland Security

Description

Fingerprint verification is vital because it provides a unique and permanent way to identify individuals. This technology is widely used in various areas like law enforcement, access control, and identity verification processes. Existing approaches for fingerprint verification tasks suffer from low accuracy due to training directly on low-quality and latent fingerprints. Therefore, this work proposes to utilize recent advancements in deep learning and computer vision to (1) enhance fingerprint image quality;(2) extract and verify that the minutiae are retained after enhancement; and (3) perform fingerprint verification tasks. Specifically, this work experiments with (1) Super-Resolution Convolutional Neural Network (SRCNN), Fast SRCNN, and Very Deep Super Resolution (VDSR) for fingerprint image enhancement;(2) Finger-Flow for minutia extraction; and (3) Siamese neural network for fingerprint verification. The experiment results indicate that among the experimented super resolution approaches, VDSR outperforms the others. Additionally, it can retain minutiae in the enhanced version and shows great potential to enhance latent fingerprints, which are less visible. Most importantly, the verification performances improve on the enhanced fingerprints versus low-resolution counterparts.

Total citations

Cited by 2

Title

Hands-On Computer Vision with Detectron2

Authors

Vung Pham

Publication date

2023/4

Pages

319

Publisher

Packt, ISBN: 9781800561625

Description

Computer vision is a crucial component of many modern businesses, including automobiles, robotics, and manufacturing, and its market is growing rapidly. This book helps you explore Detectron2, Facebook's next-gen library providing cutting-edge detection and segmentation algorithms. It’s used in research and practical projects at Facebook to support computer

vision tasks, and its models can be exported to TorchScript or ONNX for deployment. The book provides you with step-by-step guidance on using existing models in Detectron2 for computer vision tasks (object detection, instance segmentation, key-point detection, semantic detection, and panoptic segmentation). You'll get to grips with the theories and visualizations of Detectron2's architecture and learn how each module in Detectron2 works. As you advance, you'll build your practical skills by working on two real-life projects (preparing data, training models, fine-tuning models, and deployments) for object detection and instance segmentation tasks using Detectron2. Finally, you'll deploy Detectron2 models into production and develop Detectron2 applications for mobile devices. By the end of this deep learning book, you'll have gained sound theoretical knowledge and useful hands-on skills to help you solve advanced computer vision tasks using Detectron2.

Total citations

Cited by 2

Title

RDNet: Deep Learning Model for Predicting pH H20 and pHKCl from Soil Vis-NIR Spectra

Authors

Vung Pham, David C Weindorf, Tommy Dang

Publication date

2021/12/15

Conference

2021 IEEE International Conference on Big Data (Big Data)

Pages

3436-3445

Publisher

IEEE

Description

Soil properties are vital to profiling and utilizing soil resources. Conventional approaches to measurements of soil properties often involve costly, environmental-unfriendly, and time-consuming laboratory procedures. Conversely, machine learning (ML) and deep learning (DL) are gaining traction in giving rapid, non-destructive, and cost-saving alternatives to

predictions of soil properties. These ML/DL models are convenient and fast because they utilize spectral data, such as visible and near-infrared (Vis-NIR) spectra, that can be easily collected using proximal sensors for their training and prediction purposes. However, existing ML/DL approaches to this problem pose several limitations, such as having small sample sizes, needing to divide the sample data into local areas to increase accuracy, and having relatively low accuracy. Therefore, this work experiments various ML/DL methods that leverage Vis-NIR spectra collected from a rather large number of soil samples distributed all over the world to predict pHH2O and pHKCl. We then propose a DL method, called RDNet, that outperforms the other existing approaches. We also utilize visualizations to verify if the proposed model learns legitimate information from the training data.

Total citations

Cited by 2

Title

Genexplorer: Visualizing and comparing gene expression levels via differential charts

Authors

Chau Pham, Vung Pham, Tommy Dang

Publication date

2020/10/5

Book

International Symposium on Visual Computing

Pages

248-259

Publisher

Springer International Publishing

Description

This paper describes a visual interface for analyzing gene expression data generated from multiple biological samples under different controlled conditions. The tasks are to provide a comprehensive overview of thousands of genes under different states and have an intuitive way to narrow down genes with common behaviors. Our method involves using multidimensional projections and differential charts to help users analyze different data sets

via a web-based interface. Incorporating these charts and other visualization techniques into our final design makes the application accessible to genetics analysts, as demonstrated in the two use cases in plant and cancer researches. We further discuss the feedback from domain experts and the limitations of our approach to accommodate gene exploration tasks.

Total citations

Cited by 2

Title

Timematrix: Visual representation for temporal pattern detection in dynamic networks, vast 2018 mini-challenge 3

Authors

Tommy Dang, Vung V Pham

Publication date

2018/10/21

Conference

2018 IEEE Conference on Visual Analytics Science and Technology (VAST)

Pages

108-109

Publisher

IEEE

Description

This work proposes a visual analytic technique for visualizing temporal pattern detection in networks. The force-directed layout is a popular way to highlight structures of a network however it does not allow to present the dynamic features (how communities change over time). We target this problem by forcing nodes to display vertically and using the horizontal space for representing the timeline of entities. To show the connection between nodes, the edges are drawn vertically to avoid edge-crossings. This work also presents the result of applying the solution to the VAST 2018 - Mini Challenge 3 dataset, which led to the Honorable Mention: Representation of Small-Scale Temporal Patterns for the challenge.

Total citations

Cited by 2

Title

Optimizing YOLO Architectures for Optimal Road Damage Detection and Classification: A Comparative Study from YOLOv7 to YOLOv10

Authors

Vung Pham, Lan Dong Thi Ngoc, Duy-Linh Bui

Publication date

2024/12/15

Conference

2024 IEEE International Conference on Big Data (BigData)

Pages

8460-8468

Publisher

IEEE

Description

Maintaining roadway infrastructure is essential for ensuring a safe, efficient, a nd sustainable transportation system. However, manual data collection for detecting road damage is time-consuming, labor-intensive, and poses safety risks. Recent advancements in artificial intelligence, particularly deep learning, offer a promising solution for automating this process using road images. This paper presents a comprehensive workflow for road damage detection using deep learning models, focusing on optimizations for inference speed while preserving detection accuracy. Specifically, to accommodate hardware limitations, large images are cropped, and lightweight models are utilized. Additionally, an external pothole dataset is incorporated to enhance the detection of this underrepresented damage class. The proposed approach employs multiple model architectures, including a custom YOLOv7 model with Coordinate Attention layers and a Tiny YOLOv7 model, which are trained and combined to maximize detection performance. The models are further reparameterized to optimize inference efficiency. Experimental results demonstrate that the ensemble of the custom YOLOv7 model with three Coordinate Attention layers and the default Tiny YOLOv7 model achieves an F1 score of 0.7027 with an inference speed of 0.0547 seconds per image. The complete pipeline, including data preprocessing, model training, and inference scripts, is

publicly available on the project's GitHub repository, enabling reproducibility and facilitating further research.

Total citations

Cited by 1

Title

Improving road damage detection accuracy using deep learning image enhancement models

Authors

Vung Pham

Publication date

2024

Publisher

Institute for Homeland Security

Description

It is important to accurately detect and classify road damage to ensure road human safety. Current road damage detection systems suffer from detection accuracy loss due to low-quality images. This work endeavors to elevate the efficacy of road damage detection systems by leveraging state-of-the-art deep learning techniques in image enhancement. Specifically, this project experiments with different image enhancement models (including Super Resolution Residual Network or SRResNet and Super Resolution Generative Adversarial Network or SRGAN) and evaluates how they help detection models (You Only Look Once versions 7 and 9 or YOLOv7 and YOLOv9) improve their detection accuracy. The experimental results show that SRResNet helps generate super-resolution images that subsequently improve the detection performances by approximately seven times, and SRGAN, though it produces good-looking images, does not help improve the performance as much as SRResNet due to generating unrealistic patterns.

Total citations

Cited by 1

Title

Securing autonomous system in multi-domain tactical environment

Authors

Dy D Le, Vung Pham, Tommy Dang

Publication date

2020/5/22

Conference

Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II

Volume

11413

Pages

662-670

Publisher

SPIE

Description

The U.S. Army envisions fighting and winning future wars in congested and contested environments and multi-domain battles where revolutionary capabilities for the network-centric warfare (NCW) are essentially needed. NCW is characterized by the ability of geographically dispersed forces to attain a high level of shared battle-space awareness that can be exploited to achieve strategic, operational, and tactical objectives by autonomously linking people, platforms, weapons, sensors, and decision aids into a single network. Future battlefield networks will generate a massive volume of data, which can go beyond quantities. In a multi-domain battle, novel technologies for real-time decision-making, which is based on a large amount of heterogeneous as well as sparse, noisy, and ill-defined data under extremely uncertain environment, are specifically required. Additionally, humans have sometimes become completely comfortable with the information brought in by our sensing technologies. As a result, the command architecture, built on a massive web of information sources, becomes more receptive to potential catastrophic machine-human decision-making conflicts as well as vulnerable to incoming cyber threats including adversaries' deception, interruption, and obscuration, which can eventually introduce own sources of decision-making failure. In this paper, researchers present validation results of a conceptualized artificial intelligence-based visual analytics framework. The researchers' ultimate goal is to integrate the mature technology into the situation awareness technology for local commands and global logistics centers to enable an effective logistic command and control of aviation platforms and autonomous systems while being operated in an expeditionary multi-domain environment.

Total citations

Cited by 1

Title

The Effect of Ironic Process Theory on Brain Signal-Based Encryption for IoT Devices

Authors

Ahmet Furkan Aydogan, Cihan Varol, Narasimha Karpoor Shashidhar, Amar Rasheed, Van Vung Pham, Murat Karabatak

Publication date

2024/12/5

Journal

Electronics

Volume

13

Issue

23

Pages

4804

Publisher

MDPI

Description

Numerous encryption methods have been published to secure IoT devices in the last decade. Existing encryption methods still have disadvantages when it comes to securing IoT devices. On the other hand, a new encryption method using brain signals in IoT devices is gaining attention as a new solution. The encryption method based on brain signals essentially involves a hypothesis called imposed recall based on ironic process theory. The imposed recall was created with the expectation that imposing a specific choice on the subjects during the acquisition of brain signals would allow for better separation of EEG data. This paper presents experiments and approaches to prove the validity of the imposed recall hypothesis.

With the experiments, the effects of ironic process theory on brain signal-based encryption can be observed. While performing the tests, varying approaches, including Granger causality, were applied to analyze the results. The results show that the imposed recall hypothesis can successfully reconstruct EEG data. The structured signals were determined to be effective in capturing matches of brain signals on subjects at different time intervals. Thus, the imposed recall hypothesis can be used in various fields, such as authentication, questioning, and identification, by reserving brain signals to be obtained from individuals. In addition, it was reported that it is possible to acquire the ability to provide security in both devices with limited hardware, such as IoT devices or complex systems.

Total citations

Cited by 0

Title

Melanoma Classification Through Deep Ensemble Learning and Explainable AI

Authors

Wadduwage Shanika Pereraa, ABM Islam, Van Vung Pham, Min Kyung An

Publication date

2024/2

Conference

The 17th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2024)

Volume

2

Pages

263-274

Description

Melanoma is one of the most aggressive and deadliest skin cancers, leading to mortality if not detected and treated in the early stages. Artificial intelligence techniques have recently been developed to help dermatologists in the early detection of melanoma, and systems based on deep learning (DL) have been able to detect these lesions with high accuracy. However, the entire community must overcome the explainability limit to get the maximum benefit from DL

for diagnostics in the healthcare domain. Because of the black box operation's shortcomings in DL models' decisions, there is a lack of reliability and trust in the outcomes. However, Explainable Artificial Intelligence (XAI) can solve this problem by interpreting the predictions of AI systems. This paper proposes a machine learning model using ensemble learning of three state-of-the-art deep transfer Learning networks, along with an approach to ensure the reliability of the predictions by utilizing XAI techniques to explain the basis of the predictions.

Total citations

Cited by 0

Title

Improving Fingerprint Classification Accuracy Using Image Enhancement Deep Learning Models

Authors

Nikita Dalvi, Min Kyung An, Abm Rezbaul Islam, Vung Pham

Publication date

2023/12/15

Conference

2023 IEEE International Conference on Big Data (BigData)

Pages

3245-3254

Publisher

IEEE

Description

Fingerprint verification is widely used to verify individuals based on their unique fingerprint patterns. However, existing fingerprint identification systems encounter challenges while dealing with poor-quality images, smudged or marked fingerprints, and variations in finger positioning. These issues are common for latent fingerprints. This work proposes to address these issues through a novel machine-learning model employing advanced image enhancement techniques. The model aims to enhance fingerprint image quality and minimize the impact of damage using machine learning, ultimately reducing the error rate in identification systems. Specifically, this work proposes to incorporate deep learning image

enhancement techniques into the low-quality and latent fingerprints before passing them through another deep learning model to perform verification tasks. This work also provides insights as different experiments are made while applying different approaches to different real-life datasets.

Total citations

Cited by 0

Title

Comparative analysis and visualization of soil profiles at the meter spatial scale utilizing novel matrix and volume rendering techniques

Authors

Jake Gonzalez, Matthew Siebecker, Vung Pham, Cynthia Jordan, David C Weindorf, Tommy Dang

Publication date

2023/12/1

Journal

Computers and Electronics in Agriculture

Volume

215

Pages

108377

Publisher

Elsevier

Description

This research introduces a soil characterization technique involving four data visualization tools to help researchers and stakeholders interpret high dimensional soil data at the field scale. This technique involves visualizing a reduced dimensionality representation of elemental concentration and color data gathered via portable X-ray fluorescence (pXRF) spectrometer and NixPro color proximal sensors, respectively. Soil cores were collected from sites located in Lubbock and Lamb Counties, West Texas, USA. Thirteen core samples were

collected from these sites in a star pattern with readings from proximal sensors at depths ranging between 0 and 100 cm at 10 cm intervals. The dimensionality reduction techniques utilize four visualization tools to represent soil composition data through multiple user-adjustable variables (i.e., mg kg−1 elemental concentrations and soil profiles), offering more insight and control compared to a single-variable approach. Through these tools and techniques, qualitative and quantitative conclusions regarding soil characteristics (e.g., elemental concentration variation, delineation of soil horizons, changes in soil color) can be formulated from the data and used in various applications. Areas where these novel software tools can be utilized potentially include rapid contaminant mapping in soils, characterization of diagnostic soil horizons (e.g., calcic, spodic, gypsic, etc.), micronutrient distribution at a field scale for precision agricultural purposes, and pedometrics.

Title

RaCAViz: Interactive Visualizations for Rapid Carbon Assessment

Authors

Du Nguyen, Vung Pham

Publication date

2022/12/17

Conference

2022 IEEE International Conference on Big Data (Big Data)

Pages

4543-4550

Publisher

IEEE

Description

Soil organic carbon is an essential element of environmental quality assessment because carbon dioxide is the main greenhouse gas causing global warming and climate change. Trees and plants acquire a lot of carbon in their lifetimes, and soil helps keep this portion of carbon from releasing to the environment. Thus, understanding the current carbon storage in the soil

and the relationship with tree richness is important. The soil organic carbon datasets and corresponding tree information are available. However, the current visualizations for these datasets are either static visualizations that report initial findings to the public or visualizations that are too complicated to be useful for a broad audience. Therefore, this work proposes an interactive visualization solution for analyzing soil organic carbon values distribution and their relationship with tree heights and diameters at breast heights, called RaCAViz. The design of this interactive visualization is based on the analytical evaluation of visual and interaction idioms to tackle typical analysis tasks on this type of dataset. However, this paper also provides a specific use case to illustrate its usefulness in bringing different perspectives on soil organic carbon datasets to a broad audience.

Total citations

Cited by 0

Title

Development of a Forensic Toolkit for Small-Medium Size Business (SMB)

Authors

Eduardo Martinez, Cihan Varol, Narasimha Shashidhar, Van Vung Pham

Publication date

2022/6/6

Conference

2022 10th International Symposium on Digital Forensics and Security (ISDFS)

Pages

1-4

Publisher

IEEE

Description

With the increasing rate of digitalization around the world, more and more small-medium size businesses (SMBs) now depend on technology. Sources claim that more than half of SMBs experience cyberattacks, which leave certain artefacts that can be used as information to reveal the cause of the attack, if processed and analyzed correctly. This work proposes a portable toolkit that aims to assist users in the extraction and identification of forensic data

for SMBs. Through its graphical interface, users will be able to select which type of information they would like to extract, by selecting the appropriate built-in tool.

Total citations

Cited by 0

Title

HMaViz: Human-machine analytics for visual recommendation

Authors

Ngan VT Nguyen, Vung Pham, Tommy Dang

Publication date

2021/6/29

Book

Proceedings of the 12th International Conference on Advances in Information Technology

Pages

1-9

Description

Visualizations are context-specific. Understanding the context of visualizations before deciding to use them is a daunting task since users have various backgrounds, and there are thousands of available visual representations (and their variances). To this end, this paper proposes a visual analytics framework to achieve the following research goals: (1) to automatically generate a number of suitable representations for visualizing the input data and present it to users as a catalog of visualizations with different levels of abstractions and data characteristics on one/two/multi-dimensional spaces (2) to infer aspects of the user's interest based on their interactions (3) to narrow down a smaller set of visualizations that suit users analysis intention. The results of this process give our analytics system the means to better understand the user's analysis process and enable it to better provide timely recommendations.

Total citations

Cited by 0

Title

DATA VISUALIZATION DEVICE AND METHOD

Inventors

Vung PHAM, Van, C. WEINDORF, David, Tommy DANG

Publication date

2021/3/25

Patent office

WO

Patent number

WO/2021/055243

Application number

PCT/US2020/050313

Description

An apparatus and method include: receiving a data set having two or more variables; receiving a selection of at least one of the two or more variables, an abstraction level and a visual feature; automatically generating and displaying a set of visual representations of the data set on the display; receiving a change in the selected variables, selected abstraction level, the selected visual feature, or a selection from various views; determining a visual representation recommendation based on the selected variable(s), selected abstraction level and the selected visual feature, the change in the selected variables, selected abstraction level, the selected visual feature, or the selected views; and automatically updating and displaying the set of visual representations of the data set on the display based the visual representation recommendation, and the change in the selected variables, selected abstraction level, the selected visual feature, or the selected views.

Total citations

Cited by 0

Title

Integrating visual analytics and machine learning approaches for analyzing multivariate proximal sensor data

Authors

Vung Pham

Publication date

2021

Description

Sensors are pervasive in a broad range of industries. They provide efficient methods to measure and monitor properties of interest for numerous business and manufacturing operations. Recent developments in sensing technologies and their efficiency bring a massive amount of multivariate proximal sensor data. While collecting multivariate proximal sensor data becomes more accessible, analyzing them still consumes a lot of time for domain-specific analysts. Therefore, this dissertation discusses five objectives, consolidated from interviewing 102 stakeholders, for a software solution in this area. These objectives include 1) having a set of typical visualizations for chemical measurement data, 2) having an intelligent visual recommendation component that provides personalized visualizations for chemical analysts, 3) having real-life error indications for proximal sensors, 4) having machine learning components for device calibrations (eg, spectrum to concentrations), and 5) having machine learning components with an emphasis on high-level, difficult-to-measure property predictions from low-level, relatively easy-to-measure proximal sensor data.

Total citations

Cited by 0

Title

Enhanced soil profile visualization using portable X-ray fluorescence (PXRF) spectrometry.

Authors

Sun FuJun Sun FuJun, N Bakr, T Dang, Vung Pham Vung Pham, DC Weindorf, Jiang ZhuoDong Jiang ZhuoDong, Li HuaLei Li HuaLei, Wang QiuBing Wang QiuBing

Publication date

2020/1/9

Volume

358

Pages

113997

Description

Soil horizonation has traditionally relied upon morphological description, field sampling, and laboratory analysis as a means of establishing key diagnostic horizons and features within a soil profile. While this technique is useful for many soil properties, other quantifiable properties associated with elemental abundance may be visually imperceptible. Portable X-ray fluorescence (PXRF) spectrometry allows for rapid elemental data acquisition in-situ, providing more than 20 elements commonly detectable within a soil profile. In this study, a grid containing 130 cells was physically imposed upon two soil profiles in West Texas, USA. PXRF was used to sequentially scan each grid cell and acquire elemental data. The profiles were morphologically described by a panel of soil scientists, sampled, and subjected to traditional laboratory characterization. Data from the PXRF and field sampling were statistically compared, then visualized as depth functions. A data visualization tool was created to overly elemental data on top of a digital photograph of the soil profile in semitransparency. Using elemental data as a proxy, PXRF data was able to detect both argillic and calcic horizons within one soil profile and visually illustrate their extent and assist in delineating horizons in a second profile with nondescript horizons. The soil visualization tool has other features as well such as correlation analysis between elements, adjustable correlation threshold in real time, rasterized vs. smoothed data overlay, and box-plots of elemental concentrations. Summarily, this research developed a new tool for PXRF data visualization in support of soil pedon description, allowing for elemental data which may be visually imperceptible to be considered during soil profile characterization. Future advancements will seek to integrate this tool directly into the PXRF software for on-site visualization of any element specified by the analyst. Applications of such could include real-time mapping of contaminant metals in soil profiles.

Total citations

Cited by 0

Title

JavaScript Implementation of Scagnostics and Its Applications

Authors

Vung Pham, Tommy Dang

Publication date

2019

Journal

Data Science, Data Visualization, and Digital Twins

Publisher

InTech

Description

Scagnostics is a set of features that characterizes the 2D distributions in the underlying data. Various real-world applications have been using Scagnostics visual features to detect unusual bivariate data correlations. Concomitantly, many applications are required to be implemented on web platforms due to their accessibility and convenience. Therefore, this chapter discusses a recent JavaScript implementation of Scagnostics, an extension to higher dimensional data, and its applications in detecting abnormalities in bivariate and multivariate time series data. Its imple-mentation in JavaScript supports the tremendous demand for visual features in the web environment. Likewise, its higher dimensional implementations allow gener-ating Scagnostics features for the rapidly growing multivariate data. Finally, conventional ScagnosticsJS computations involve time-consuming algorithms, and they are sensitive to slight changes in the underlying data. Therefore, this chapter also discusses a recent attempt to tackle these issues using machine learning to estimate the Scagnostics scores.

Total citations

Cited by 0

Title

Interactive Multidimensional Visual Analytics for Earth’s Mantle Convection

Authors

Jansen Wong, Vung Pham, Tommy Dang

Description

This project creates an interactive visualization of the Earth’s mantle convection. We first process the data using Python; then, the processed data is visualized in the browser using JavaScript. The application uses parallel coordinates, a correlogram, a line graph, cross-sections and volume renderers for visual comparison and qualitative analysis of the Earth’s mantle convection.

Total citations

Cited by 0

Title

ScagnosticsJS: Extended Scatterplot Visual Features for the Web Supplementary Document

Authors

V Pham, T Dang

Description

Figure 2: Scagnostics measures and their targeting patterns projected using two and three PCA principal components in panel (a) and panel (b) correspondingly. three-dimension (3D), and higher dimension (nD) using our Scagnostics JavaScript implementations. The color scale is a linear one ranging from blue (for the minimum score of 0.0) to red (for the maximum score of 1.0). It is observable that the measurements are all close to red (ie, their corresponding values are closer to 1.0) for the patterns that they target (the diagonal cells with borders).

Total citations

Cited by 0

Title

HPCViz: Monitoring Health Status of High Performance Computing Systems

Authors

Tommy Dang, Vinh T Nguyen, Vung Pham, Ghazanfar Ali, Yong Chen

Description

This paper introduces HPCViz, a visual analytic tool for tracking and monitoring high-performance computing (HPC) system events through a RESTful interface. The goals of this tool are: 1) to monitor a set of system events from multiple hosts and racks in real-time statistics, 2) to support system administrators in alarming and detecting unusual signaturebased patterns exhibited by health records of hosts in a complex system, and 3) to help in performing system troubleshooting and maintenance with a visual layout for both computing resource allocation and health monitoring map that represent the actual system. A case study was conducted on a medium-scale, Redfishenabled production HPC system with a total of 10 racks and 467 hosts. The result of the case study shows that the visualization tool offers excellent support for system administrators and analysts to profile and observe system behavior and further identify the traces of issues occurred.

Total citations

Cited by 0