Selected Ph.D. Thesis Abstracts

This Ph.D thesis abstracts section presents theses defended in 2021 and 2022. These submissions cover a range of research topics and themes under intelligent informatics, such as prophylactic treatments to misinformation and disinformation, fatal disease detection using hybrid deep learning, geriatric care monitoring, graph model, interactive visualization, machine learning-assisted corpus exploration, large scale data exploration, remote patient monitoring, representation learning, secure content delivery in IoT, soft biometrics-based person retrieval, and safety solutions in smart cities.

AN INTERDISCIPLINARY ASSESSMENT OF THE PROPHYLACTIC EDUCATIONAL TREATMENTS TO MISINFORMATION AND DISINFORMATION

Kevin Matthe Caramancion
kcaramancion@albany.edu
University at Albany, State University of New York, USA

MISINFORMATION and Disinformation, both types of Information Disorder in the cyber world, operate on a systemic level. Several factors enabling their persistence, including laws, policies, and technological mediators, have been investigated in the literature. Cybersecurity frameworks and guidelines specify that the target victims, as part of the human factors, hold a degree of responsibility for the persistence of any threat. This dissertation re-shifts the lens to the intended targets of the falsehood attacks, the information consumers. This factor of consideration was investigated through its three-part, multi-method research phases.

The first phase, an interdisciplinary qualitative exploration of the different styles and techniques of preventive/prophylactic educational campaigns and theories that can be effectively used in raising awareness of the information consumers by employing a focus group of experts, revealed that ALA's CRAAP (Currency, Relevance, Authority, Accuracy, and Purpose) has the advantage and sustained its classical value over its counterpart prebunking (from the theory of inoculation). On the other hand, the second phase, where an outright comparison of the existing and suggested preventive treatments from the previous phase through quantifications of theories and actual user experimentations, revealed that users under the CRAAP treatment displayed greater detection accuracy (DV1) than the users on prebunking treatment. However, users on prebunking treatment resulted in a faster assessment time (DV2) compared to the users under the CRAAP group. Finally, in the third and final phase, this dissertation redesigned and improved the prevailing mis/disinformation predictive models by including the educational conditioning of users, as a factor, in forecasting, in multivariate format, their cyber risk from these attacks of deception.

Beyond policy implications, this project's significance includes contributions to the improvement of methods in scientific inquiry within the domain of risk modeling by combining self-reported data with digital trace data. Furthermore, this dissertation takes pride in its interdisciplinary design, considering the best practices of the granular fields involved to enable a more thorough investigation and probe through integration.

FATAL DISEASES DETECTION FROM ECG SIGNALS AND MRI IMAGES USING HYBRID DEEP LEARNING MODELS

Hari Mohan Rai
harimohanrai@gmail.com
Department of Electronics & Communication Engineering
Dronacharya Group of Institutions, Greater Noida, Uttar Pradesh, India

AUTOMATED and computerized detection and diagnosis of critical diseases or abnormalities from the human body with the highest accuracy is a very vital task for medical experts, researchers, and scientists. The work is divided into two segments: arrhythmias detection from Electrocardiogram (ECG) signals and tumor detection from MRI images. The ECG signal is obtained by recording the electrical activity of the human heart, and it is a noninvasive method employed as a principal diagnostic tool for the detection of cardiac arrhythmias. We know that the major disadvantage of the first system is that it is not very effective and is not suited for big datasets. Furthermore, the performance of the first system is dependent on the manual feature extraction method.

The six types of arrhythmias detection from ECG dataset has been employed using hybrid deep learning models (CNN-LSTM). To verify the model's performance, 123,998 ECG beats from the MIT-BIH arrhythmias database (MITDB) and the PTB diagnostic database (PTBDB) have been utilized. Because both datasets are extremely unbalanced, three data balancing methods have been employed to resample and balance the dataset in order to enhance the accuracy of the minority class classifications in both datasets. Using the SMOTE-Tomek link sampled dataset, the ensemble method is able to achieve an overall accuracy of 99.10 percent on a test dataset of 24,800 observations. The proposed hybrid deep learning models on big datasets offer very high accuracy for the detection of cardiac arrhythmias from ECG signals. Another hybrid deep learning model, UnetResNext-50, has been proposed for the detection of brain tumors from magnetic resonance imaging (MRI) images, which is one of the most deadly diseases. It is primarily the merging of the features of two distinct deep learning models, UNet and ResNext-50, with some modifications to the layers in the hybrid deep learning model that has been proposed.

The proposed model with the ResNext-50 backbone omits associations that take care of the degradation issue of deep

CNN models with more layers while improving pixel quality toward the vanilla Unet decoder. Total magnetic resonance imaging (MRI) dataset consisting of 3929 MR images, comprising 1373 images with tumors and 2556 images of non-tumorous kind (without tumors). The MRI dataset is primarily preprocessed by resizing, cropping, pixel normalization, and data splitting techniques before applying for segmentation and detection tasks. Also, the dataset used to train the proposed model is insufficient, therefore 12 different data augmentation (DA) functions are employed to generate a big training dataset (42,432) that can be used to train the DL models effectively. The six types of performance measurement metrics used to evaluate model proficiency are the Intersection over union or Jaccard index, F1-score, DICE score, precision, accuracy, and recall. The proposed model (UnetResNext-50) performance has been compared to two other deep learning models, Vanilla U-Net and UnetResNet-50, utilizing six types of performance assessment metrics to validate the system efficiency and accuracy. Post-processing techniques that use DICE and intersection over union (IoU) values are now being used to enhance tumor segmentation visibility. The developed hybrid deep learning model has great accuracy and accuracy for tumor identification, and it has 99.7score.

FREELANCING GERIATRIC CARE MONITORING SYSTEM IN AUSTRALIA

Hamid Ali
hamydaly@gmail.com
University of Southern Queensland, Australia

Due to the advancement in medical science, extreme decline has been seen in the death rate. Life expectancy has increased considerably, and people are enjoying a long life. Because of the reduction in the mortality rate, 20 percent of the world's population is expected to be 60 or older by 2050. It is quite common to transfer elderly people to Residential Aged Care Facilities (RACFs) from in-home care in developed countries like Australia. But due to the increasing elderly people and shortage of Personal Care Assistants (PCAs), RACFs are not enough to provide care and services. With the RACFs, elderly people lose their autonomy, independence and social interaction, to name a few things. That's why most Australians prefer to live in their homes and are frequently visited by the PCAs. AI-enabled freelancing strategy can help the geriatric care monitoring system in managing the demand and supply of PCAs in Australia.

In this paper, we propose AI-enabled freelancing strategy to help the geriatric care monitoring system in managing the PCAs. The goal is to propose a cheaper, more inclusive, and novice idea of integrating AI with freelancing strategy. Traditional research has been unable to solve this disproportionate issue of increasing elderly people and decreasing PCAs. Traditional research talks about the remote patient monitoring but that still have so many issues up to date. This novice idea will not only track but also predict the vital signs (body temperature, pulse rate, respiration rate, blood pressure etc) and send the signals to PCAs when help is needed. Vital signs are useful in detecting or monitoring medical problems at home. In the previous models, PCAs are bound in their duties, and they pay a visit to the elderly people without even being asked for. Since the Australian Government uses pay-as-you-go strategy, which means younger people are paying for the aged-care through taxation, it's going to be difficult for the government to manage in the future with less young generation and an increasing older generation.

There are many technologies that can be used but they are too expensive to implement, or the solutions are not concrete enough to solve the issue of aging people. Different solutions like ambient and wearable sensing, deploying of VR and implementation of 'smart shoe insoles' are too costly to implement. Previous research has given qualitative solutions while this research, which is of an exploratory nature, proposed a concrete solution to tackle the current problems and add to previous literature. With this freelance work, it will take the burden off full-time PCAs in the future. Similarly the intake of aged-care workers has been increasing, which will help to meet the growing demand of PCAs, as attracting and retaining workers is a big problem. The main purpose of clustered and in-home care is to support elderly people and giving the residents more autonomy in their routines and more flexibility when it comes to activities and outdoor access. So, this is a more flexible model, and it is going to help the Australian aged-care sector in the coming decades.

GRAPH MODEL FOR SCHEMA AND DATA MAPPING

Sonal Tuteja
sonalt9@gmail.com
Jawaharlal Nehru University
New Delhi, India

Graph model has emerged as a data modeling technique for large-scale applications with heterogeneous, schemaless, and relationship-centric data. It can help unveil relationships otherwise hidden in heterogeneous data sources. The flexibility and schemaless nature of graph models have led to various modeling techniques to map data from several data models into a graph model. However, the unification of data from heterogeneous source models into a graph model has not received much attention. Furthermore, designing graph models based on queries to be addressed has not been explored much. In addition, the role of data mapping techniques in a graph model for query performance has not been considered.

Addressing these research gaps, we propose a framework for unifying heterogeneous data sources into a graph model. We also analyze and compare the unified graph's query performance, scalability, and database size with heterogeneous source data models. We design various graph models for an e-commerce application incorporating queries to be executed and compare their performance with the baseline graph model. We define different data mapping techniques for graph models and verify their equivalence with the source graph model.

We observed that the graph model outperformed the relational and ontology models in all performance measures, except for aggregation queries. In query-driven graph models,

incorporating new nodes and edges improved the query performance of selection, projection, path-traversal operations, and their combinations. The designed data mapping techniques can be used in graph models for creating different relationships. Thus, our thesis designs and develops frameworks and techniques for improving a graph model's performance.

INTERACTIVE VISUALIZATION FOR INTERPRETABLE MACHINE LEARNING

Dennis Collaris
d.a.c.collaris@tue.nl
Eindhoven University of Technology Netherlands

MACHINE learning has firmly established itself as a valuable and ubiquitous technique in commercial applications. But these models are often complex and difficult to understand. Understanding models is particularly important in high-impact domains such as credit, employment, and housing, where the decisions made using machine learning impact the lives of real people. The field of eXplainable Artificial Intelligence (XAI) aims to help experts understand complex machine learning models. In recent years, various techniques have been proposed to open up the black box of machine learning. However, because interpretability is an inherently subjective concept it remains challenging to define what a good explanation is. We argue we should actively involve data scientists in the process of generating explanations and leverage their expertise in the domain and machine learning. Interactive visualization provides an excellent opportunity to both involve and empower experts.

In this dissertation, we explore interactive visualization for machine learning interpretation from different perspectives, ranging from local explanation of single predictions to global explanation of the entire model. We first introduce ExplainExplore: an interactive explanation system to explore explanations of individual predictions (i.e., local). For each explanation, it provides context by presenting similar predictions, and showing the impact of small input perturbations. We recognize many different explanations may exist that are all equally valid and useful using traditional evaluation methods. Hence, we leverage the domain knowledge of the data scientist to determine which of these fit their preference. To ensure these contributions can be broadly applied, we introduce a software library that enables interoperability with a wide range of different languages, toolkits, and enterprise software. Next, we propose the Contribution-Value plot as a new elementary building block for interpretability visualization, showing how feature contribution changes for different feature values. It provides a perspective in between local and global, as the model behavior is shown for all instances, but visualized on a per-feature basis. In a quantitative online survey with 22 participants, we show our visualization increases correctness, confidence, and reduces the time needed to obtain an insight compared to previous techniques.

This work highlighted that a small difference in feature importance techniques can result in a large difference in interpretation, and warranted a follow-up human computer interaction contribution to characterize the data scientists' mental model of explanations to explore the differences between existing techniques. Finally, we introduce StrategyAtlas: a visual analytics approach to enable a global understanding of complex machine learning models through the identification and interpretation of different model strategies. These model strategies are identified in our projection-based StrategyMap visualization. Data scientists can ascertain the validity of these strategies through analyzing feature values and contributions using heat maps, density plots, and decision tree abstractions. As computing the local feature importance values for an entire dataset is computationally expensive, we complement this work with an algorithmic contribution called LEMON to improve the faithfulness of explanation results, enabling significantly sped up computations of StrategyMap projections.

MACHINE-LEARNING-ASSISTED CORPUS EXPLORATION AND VISUALISATION

Tim Repke
repke@mcc-berlin.net
Information Systems Group, Hasso Plattner Institute at the University of Potsdam, Germany
https://hpi.de/naumann/people/tim-repke.html

TEXT collections, such as corpora of books, research articles, news, or business documents are an important resource for knowledge discovery. Exploring large document collections by hand is a cumbersome but necessary task to gain new insights and relevant information. Our digitized society allows us to utilize algorithms to support the information seeking process, for example, with the help of retrieval or recommender systems. However, these systems only provide selective views of the data and require prior knowledge to issue meaningful queries and asses a system's response. The advancements of machine learning allow us to reduce this gap and better assist the information seeking process. For example, instead of sighting countless business documents by hand, journalists and investigators can employ natural language processing techniques such as named entity recognition.

Although this greatly improves the capabilities of a data exploration platform, the wealth of information is still overwhelming. An overview of the entirety of a dataset in the form of a two-dimensional map-like visualization may help to circumvent this issue. Such overviews enable novel interaction paradigms for users, which are similar to the exploration of digital geographical maps. In particular, they can provide valuable context by indicating how a piece of information fits into the bigger picture. This thesis proposes algorithms that appropriately preprocess heterogeneous documents and compute the layout for datasets of all kinds. Traditionally, given high-dimensional semantic representations of the data, so-called dimensionality reduction algorithms are used to compute a layout of the data on a two-dimensional canvas. In this thesis, we focus on text corpora and go beyond only projecting the inherent semantic structure itself. Therefore, we propose three dimensionality reduction approaches that incorporate additional information into the layout process: (1) a multi-objective dimensionality reduction algorithm to jointly

visualize semantic information with inherent network information derived from the underlying data; (2) a comparison of initialization strategies for different dimensionality reduction algorithms to generate a series of layouts for corpora that grow and evolve over time; (3) and an algorithm that updates existing layouts by incorporating user feedback provided by pointwise drag-and-drop edits. In the scope of this thesis, we also developed system prototypes to demonstrate the proposed technologies, including pre-processing and layout of the data and presentation in interactive user interfaces.

MINIMIZING USER EFFORT IN LARGE SCALE EXAMPLE-DRIVEN DATA EXPLORATION

Xiaoyu Ge
xiaoyu@cs.pitt.edu
Department of Computer Science University of Pittsburgh, USA
https://d-scholarship.pitt.edu/41742

THE ever-increasing supply of data is bringing renewed attention to data exploration, a technique that serves as a key ingredient in a widely diverse set of discovery-oriented applications, including scientific computing, financial analysis, and evidence-based medicine. One major challenge for those discovery-oriented applications is the need to extract useful pieces of knowledge from data while requiring little to no specification of the information that is being searched for. As the traditional searching and data mining techniques fall short in meeting such challenging demands, data exploration techniques aimed at intelligently assisting users in constructing precise exploratory queries have recently generated a lot of interest in both the academic and industrial communities and have led to the development of a variety of semi-automatic data exploration approaches. Among such approaches, Example-driven Exploration is rapidly becoming an attractive choice for exploratory query formulation since it attempts to minimize the amount of prior knowledge required from the user to form an accurate exploratory query.

This dissertation focuses on interactive Example-driven Exploration, which steers the user toward discovering all data objects relevant to the users' exploration based on their feedback on a small set of examples (i.e., data objects selected from the underlying dataset). Interactive Example-driven Exploration is especially beneficial for non-expert users as it leverages human-in-the-loop paradigms and enables them to circumvent query languages by assigning relevancy to the presented examples and leveraging them as proxies for the intended exploratory analysis. However, existing interactive Example-driven Exploration systems fall short of supporting the need to perform complex explorations for data that are large, unstructured, or high-dimensional. To overcome these challenges, in this dissertation, we have developed novel methods that facilitate the Example-driven Exploration paradigm in four different areas: data reduction, example selection, data indexing, and result refinement, which help to support largescale complex discovery-oriented applications.

The novelty of our approach is anchored on leveraging active learning and query optimization techniques. The prior enables semi-automatic exploration and reduces manual user effort. The latter helps to reduce the potential exploration space for the exploration system. Together they strike a balance between maximizing accuracy and minimizing user effort in providing feedback while enabling interactive performance on the system level for exploration tasks with arbitrary, large sized datasets. Furthermore, our proposed approach extends the exploration beyond the traditional structured data by supporting a variety of high-dimensional unstructured data and enables the refinement of results, which prevents the results from being overwhelming to the user when the exploration task is associated with too many relevant data objects. To affirm the effectiveness of our proposed models, techniques, and algorithms, we implemented multiple prototype systems and evaluated them using real-world datasets. Some of them have also been incorporated into domain-specific analytic tools. Our comprehensive evaluations have shown that our exploration methods help to reduce users' manual effort by up to 9x while achieving the same accuracy as the state-of-the-art alternatives. Furthermore, our data reduction and result refinement methods significantly reduced the system run-time and achieved a speedup of up to 159x when exploring large and complex datasets.

REMOTE PATIENT MONITORING SYSTEM USING ARTIFICIAL INTELLIGENCE

Thanveer Shaik
Thanveer.Shaik@usq.edu.au
University of Southern Queensland, Toowoomba, Australia

HEALTHCARE applications are vastly dependent on artificial intelligence (AI) methodologies to enable forecasting capability and provide utmost care to patients in hospital as well as remotely. Especially in psychiatric care with acute mental illness and depressed suicidal tendency patients, the goal is to provide a safe and therapeutic environment to both patients and medical staff by avoiding the physical violence caused by aggressive and agitated patients. This could be achieved by monitoring patients continuously to detect their movements and vital signs such as heart rate, respiratory rate, and breathing. The frequency of manually recorded patient reports is limited due to staff availability and number of patients in a hospital. Remote patient monitoring (RPM) could enable the continuous monitoring of acutely ill patients in psychiatric care. Empowering the RPM strategy with AI methods could transform healthcare monitoring by predicting patients' future vital signs and also classify their body movements.

In this study, the aim is to propose an AI enabled RPM system with non-invasive technology. Traditional RPM systems with invasive technology such as electrocardiography (ECG) and photoplethysmography (PSG) touch patients' skin to record their data and could cause inconvenience, limiting their daily actions. We propose a non-invasive technology radio frequency identification (RFID) based on near-field coherent sensing (NCS) without touching patients' body and allow their daily activities. RFID passive tags will need to be arranged at different areas of body such as chest area, abdomen, and limbs to record the vital signs and their body

motion. The passive tags data will be retrieved via RFID reader-antennas to a computer to process and retrieve patients' vital signs. Advanced AI methodologies such as reinforcement learning, explainable AI (ExAI) and federated learning will be adopted for adaptive learning of patients' behavior, to enhance interpretability and transparency in AI model results for decision-making, and enable personalized monitoring with patient privacy, respectively. The adaptive learning will consider each patient as an individual learning agent in a hospital environment, attempting to achieve maximum rewards by following the designed policy of staying safe clinically.

In ExAI, AI models will be interpreted based on their weights and also estimate Shapley values to extract feature importance, and model behavior. Personalized monitoring will be achieved by adopting a federated learning approach in which each patient data will be monitored individually using a local AI model and pass only the model parameters or predictions to build a robust global model. The proposed AI enabled RPM system would monitor patients' behavior adaptively while forecasting their future vital signs and classifying their physical activities. This research contributes a novel patient monitoring system that learns patient behavior, and assists clinicians with a decision support system that makes timely interventions and avoids acute disturbances in psychiatric care.

REPRESENTATION LEARNING FOR TEXTS AND GRAPHS: A UNIFIED PERSPECTIVE ON EFFICIENCY, MULTIMODALITY, AND ADAPTABILITY

Lukas Galke
Lukas.Galke@mpi.nl
Kiel University, Germany Max Planck Institute for Psycholinguistics, Netherlands

FUELED by deep learning, natural language processing is becoming increasingly influential. Meanwhile, graph representation learning shows how to process graph data effectively. However, the size of language models and large, evolving graphs becomes increasingly challenging. The immense computing power and GPU memory requirements make it difficult for small companies and research labs to participate. Thus, this thesis aims to find efficient text and graph representation learning methods that can adapt to new data. This thesis is situated between text and graph representation learning and investigates selected connections.

First, we introduce matrix embeddings as an efficient text representation sensitive to word order. After self-supervised pretraining, the matrix product acts as sentence encoding for downstream tasks. Experiments with ten linguistic probing tasks, 11 supervised, and five unsupervised downstream tasks reveal that vector and matrix embeddings have complementary strengths and that a jointly trained hybrid model outperforms both. Second, a popular pre-trained language model, BERT, is distilled into matrix embeddings. To this end, we extend matrix embeddings with a bidirectional component and equip them with a strategy to encode sentence pairs. The results on the GLUE benchmark show that these models are competitive with other recent contextualized language models while being more efficient in time and space. Third, we compare three model types for text classification: bag-of-words, sequence-, and graph-based models. Experiments on five datasets show that, surprisingly, a wide multilayer perceptron on top of a bag-of-words representation is competitive with recent graph-based approaches, questioning the necessity of graphs synthesized from the text. Pretrained Transformer-based sequence models perform best but come with high computational costs.

Fourth, we investigate the connection between text and graph data in document-based recommendation systems for citations and subject labels. Experiments on six datasets show that the title as side information improves the performance of autoencoder models. We confirm this result under different experimental conditions: the number of all possible items and the fraction of already-present items per document. We find that the meaning of item co-occurrence is crucial for the choice of input modalities and an appropriate model. Fifth, we introduce a generic framework for lifelong learning on evolving graphs in which new nodes, edges, and classes appear over time. The task is to classify nodes and detect new classes based on textual and graph information. We experiment with five representative graph neural network models and three datasets based on scholarly articles: two citation graphs and one collaboration graph.

The results show that by reusing previous parameters in incremental training, it is possible to employ smaller history sizes with only a slight decrease in accuracy compared to training with complete history. Furthermore, weighting the binary cross-entropy loss function is essential for automatically detecting newly emerging classes. This work opens up new opportunities for efficient text and graph representation learning. It shows how recommender systems can exploit textual side information and lays the foundation for lifelong and open-world learning in evolving graphs with text-attributed nodes.

SECURE CONTENT DELIVERY IN TWO-TIER CACHE-AIDED SATELLITE INTERNET OF THINGS NETWORKS

Quynh Tu Ngo
T.Ngo@latrobe.edu.au
La Trobe University, Melbourne, Australia

INTEGRATING satellites into terrestrial communication systems has been identified as a critical solution in the new era of 6G to enable global connectivity for the Internet of Things (IoT). With better service reliability and coverage, satellite communications have been an effective complementary component for IoT in addition to cellular networks. However, embedding satellites into IoT networks poses challenges in excessive service delays, expensive cost, and security issues when relating space communications. A content delivered to users is forwarded from an Internet-connected gateway through satellites and ground stations, which extends the serving time in addition to very pricey and often limited satellite bandwidth. A large-scale wireless network created when embedding satellites into IoT is exposed to more security risks due to satellite broadcast nature and wide coverage.

To tackle these challenges, this thesis first studies edge caching techniques in satellite-terrestrial network (STN) to guarantee latency and cost of services, and then physical layer security (PLS) techniques are exploited in cache-enabled STN to ensure data confidentiality. We begin by examining a two-tier cache-enabled STN model with a focus on the content delivery latency analysis in terms of the successful delivery probability (SDP). We consider enabling caching capacity in combination with fullduplex transmission at both satellite and ground station to shorten service delivery time and reduce in-network traffic. We derive a closed-form expression for the SDP given consideration to the requested content distributions, realistic channel statistics, and three commonly studied caching configurations.

Based on the derived results, we investigate the system SDP performance under different transmission modes and network settings. The network capacity in terms of maximum number of supportable users, satellite bandwidth and energy consumption are also studied. Then we design an SDP maximization-based cache placement strategy subjected to caching capacity constraints at satellite and ground station. We then secure the transmission of the two-tier cacheenabled STN by exploiting the intelligent reflecting surfaces (IRS). A novel two-hop secure content delivery scheme is proposed, and the PLS performance is investigated in terms of the secure transmission probability. Closed-form expressions for the connection probability and secrecy probability over two cascaded fading channels, i.e., the Rayleigh–Rayleigh and the Rayleigh–Shadowed-Rician fading channels are derived. Based upon the derived results, we form the system secure transmission probability, which we aim to maximize when jointly designing the transmission rates and caching probability at satellite and ground station.

To further explore the potential of IRS enhancing PLS systems, we study a hybrid IRS-assisted secure multiuser multiple-input single-output STN and analyze system performance in terms of worst-case secrecy sum-rate. A robust and secure beamforming design problem is formulated for satellite as well as hybrid IRS under practical outdated channel state information and power consumption models. We leverage deep reinforcement learning (DRL) to solve the problem by proposing a fast DRL algorithm, namely deep post-decision state–deterministic policy gradient (DPDS-DPG). DPDS-DPG exploits prior known system dynamics by integrating the PDS concept into the traditional deep DPG (DDPG) algorithm, resulting in faster learning convergence. Simulation results show better learning efficiency of DPDSDPG than DDPG with comparable achievable system secrecy rate, demonstrating the performance gains of employing hybrid IRS over conventional passive IRS to support secure communications.

SOFT BIOMETRICS-BASED PERSON RETRIEVAL FROM UNCONSTRAINED SURVEILLANCE VIDEO

Hirenkumar Jagdishchandra Galiyawala
hirenkumar.g@ahduni.edu
Ahmedabad University, Gujarat, India
http://hdl.handle.net/10603/429586

In today's security world, one interesting direction is to search and locate a person of interest from unconstrained surveillance videos using a textual description. An automated computer vision system that translates a description and retrieves the person would greatly assist. A description, e.g., a tall female with a grey short sleeve t-shirt and white jeans with a handbag, is a prime entity for retrieval. It is because the description has personal attributes (soft biometrics) like height, gender, clothing color, and clothing type. This thesis considerably contributes to developing state-of-the-art approaches for retrieving person(s) from videos using a soft biometrics-based textual description. The contributions are as follows; (1) The development of cascade filter-based approaches incorporates adaptive torso patch extraction, better height estimation, and Intersection-over-Union (IoU)-based bounding box regression, (2) The design and implementation of an attributes' weighting module for a person ranking-based approach, (3) The creation of a dataset for person attribute recognition, (4) The development of a multi-attribute learning-based model with fewer parameters which predicts all attributes using a single model, (5) The design and development of a state of the art end-to-end person retrieval approach.

The initial approach uses semantic segmentation for precise feet and head point extraction for better height estimation using the camera calibration approach. It uses only three attributes, i.e., height, gender, and torso color, where gender and color models are trained using AlexNet. This approach achieves an average IoU of 0.363, a performance of 0.522 with IoU $\geq$ 0.4, and a True Positive Rate (TPR) of 54.12%. It was improved by introducing adaptive torso patch extraction, bounding box regression, better data augmentation for color and gender models, and DenseNet-169 classification models. Person retrieval is done using height, torso type, torso color I, torso color II, and gender. This approach performs well for the person with torso type "no sleeve" as chances of unwanted pixels from the human body are minimized to improve the classification accuracy. It achieves an average IoU of 0.569, 0.746 with IoU $\geq$ 0.4, and a TPR of 76.21%.

Further improvement is made by utilizing the contribution of each attribute and a ranking strategy. This strategy performs well for a person with partial occlusion, and the resulting approach outperforms previous approaches and achieves state-of-the-art performance. This approach achieves an average IoU of 0.602, 0.808 with IoU $\geq$ 0.4, and TPR of 82.14%. Introducing a single attribute-recognizable model enables a solution toward an end-to-end person retrieval system. Considering features' correlation, a better dataset and attention mechanism with focal loss proved best for attribute recognition. It outperforms existing techniques by utilizing only five attributes. This approach achieves an average IoU of 0.667, 0.856 with IoU $\geq$ 0.4, and a TPR of 85.30%. Findings from this dissertation contribute to achieving the best performance in person retrieval from unconstrained surveillance video using soft biometrics. The work is supported by the BRNS, Govt. of India grant to supervisor Mehul S Raval.

TOWARDS SUSTAINABILITY AND SAFETY SOLUTIONS IN A SMART CITY

Federica Rollo
federica.rollo@unimore.it
Enzo Ferrari Engineering Department University of Modena and Reggio Emilia Modena, Italy

A smart city is a place where technology is exploited to help Public Administrations make decisions. Technology can contribute to managing multiple aspects of everyday life, offering more reliable services to citizens and improving the quality of life. However, technology alone is not enough to make a Smart City; suitable methods are needed to collect data generated by technology, analyze and manage them in such a way as to produce useful information. The thesis focuses on two aspects of Smart Cities: sustainability and safety. The first aspect is addressed by studying the impact of vehicular traffic on air quality through the development of traffic and air quality sensor networks and the implementation of a chain of simulation models.

This work is part of the European Union co-financed TRAFAIR project, which aims to monitor in real-time and predict air quality on an urban scale in 6 cities, including the Italian city of Modena. The project requires the management of a large amount of heterogeneous data and its integration on a scalable platform. To manage spatio-temporal data, the Smart City data platform has been implemented as a PostgreSQL database (60+ tables and 435 GB of data in 2 years - only for Modena). The simulation models used in the project to reconstruct traffic flow and forecast air quality in the cities are based on sensor data. Since sensors are prone to errors, a data cleaning process is needed to ensure reliable results of the simulation. After studying the sensor data distribution and the correlation among sensors, several anomaly detection techniques have been implemented. A novel approach employing a flow-speed correlation filter, STL decomposition and IQR has been developed for traffic sensor data as well as an innovative ensemble method for air quality sensors considering the influence between pollutant measurements. The evaluation of real-world data of Modena demonstrated the efficiency and effectiveness of these techniques.

In the thesis, the safety aspect is examined by the development of a crime analysis project which aims at generating timely and pertinent information for crime reduction, prevention, and evaluation. Due to the lack of official data, this project exploits online news articles. The goal is to categorize news articles based on crime category, geolocate crime events, detect the date of the event, and identify other relevant features (e.g., what has been stolen during the theft). A novel framework has been developed for the analysis of news articles, the extraction of semantic information by using NLP techniques, and the connection of entities to Linked Data. The emerging technology of word embeddings has been employed for text categorization. News articles referring to the same event have been identified through the application of cosine similarity. Finally, a dashboard has been developed to show the geolocated events and provide statistics and annual reports. The framework allowed the production of the Italian Crime News dataset (available online), collecting 15,000+ news articles. The impact and scalability of such a framework has been evaluated on two online newspapers of Modena. This is the first framework that, starting from Italian news articles, provides analyses of crimes and makes them available through a visualization tool.