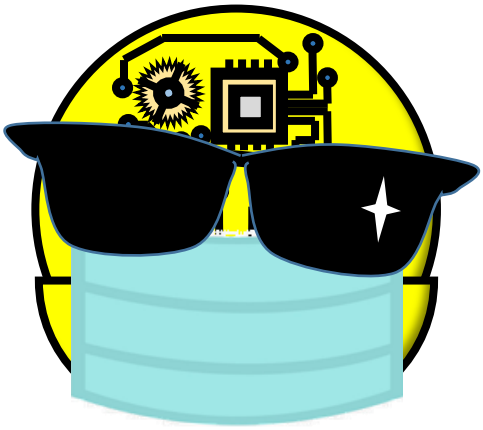
Proceedings of

Machine Intelligence Day

December 4 2020



Hosted by
Seidenberg School of Computer Science and Information Systems, Pace University
New York, New York

edited by Sung-Hyuk Cha
Paul D. Benjamin

Proceedings of Graph Theory Day 70

Published by Pace University Press

Available online
http://csis.pace.edu/∼scha/MID2020

Table of Contents

Preface

We are very pleased to have the opportunity to organize the second Machine Intelligence Day 2020. Machine Intelligence Day is an annual New York based conference hosted by Seidenberg School of Computer Science and Information Systems at Pace University. It occupies a unique place among conferences, presenting both new research and exceptional student papers, providing opportunities for both faculty and student participation. The purpose of Machine Intelligence Days is to provide a learning and sharing experience on recent developments in Artificial Intelligence, Computer Vision, Data Mining, Machine Learning, and Pattern Recognition. The conference is welcoming to a range of participants, open to both researchers in the field and students. While experts give talks, they are targeted at audiences in general computer science with an eye dedicated towards students. We have strived to publish well-written abstracts that present important original research results and/or open problems relevant to Machine Intelligence.

Two eminent invited speakers, Professor James Geller from the New Jersey Institute of Technology (NJIT) and Mr. Liam Fistos from M Science, have contributed to the conference. We are grateful to them.

Machine Intelligence Day 2020 is held online due to COVID-19. Poster session is canceled, but the recorded video presentation competition is newly organized. We received seventeen abstracts and their pre-recorded video presentations for the competition. We would like to express our gratitude to all the contributors and participants. Finally, we hope that you will benefit from this conference and its proceedings.

S.-H. Cha and D. P. Benjamin

Mining Concepts for a COVID Interface Terminology for Annotation of EHRs

V. Keloth[1], S. Zhou[1], L. Lindemann[2], G. Elhanan[3], A. Einstein[4], J Geller[1], and Y. Perl[1]
[1]Computer Science Department, NJ Institute of Technology (NJIT), Newark, NJ, USA
[2]Yale Center for Medical Informatics, Yale University, New Haven, CT, USA
[3]Renown Institute for Health Innovation, Desert Research Institute, Reno, NV, USA
[4]Department of Medicine, Cardiology Div., Columbia University, New York, NY, USA
{vk396,sz23,geller,perl}@njit.edu, gelhanan@gmail.com, luke.lindemann@gmail.com,
ae2214@cumc.columbia.edu

Due to the COVID-19 pandemic, every country in the world that is reporting trustworthy data has witnessed unprecedented and overwhelming situations for its healthcare sector. Furthermore, the emergence and global spread of new infectious diseases are highly likely to continue in the future. The lack of information on the presentations, signs, and symptoms of COVID-19 has increased the number of casualties to a high extent. The EHRs of US hospitals have ingested huge volumes of relevant, up-to-date data about patients, but the lack of a proper system to annotate this data has greatly reduced its usefulness. We propose to design a COVID interface terminology for the annotation of EHR notes of COVID-19 patients. The initial version of this interface terminology was created by integrating COVID concepts from existing ontologies such as the Coronavirus Infectious Disease Ontology (CIDO) [1]. A second major source for this work was the SNOMED CT [2], which provided necessary concepts that are not directly related to COVID.

Further enrichment of the interface terminology is performed by mining high granularity concepts from EHRs, because such concepts are usually not present in the existing reference terminologies. We use the techniques of concatenation and anchoring iteratively to extract high granularity phrases from the clinical text. In addition to increasing the conceptual base of the COVID interface terminology, this will also help in generating training data for large scale concept mining using machine learning techniques. Having the annotated clinical notes of COVID-19 patients available will help in speeding up research in this field.

References

[1] He, Y., Yu, H., Ong, E., Wang, Y., Liu, Y., Huffman, A., Huang, H., Beverley, J., Hur, J., Yang, X., Chen, L., Omenn, G., Athey, B., and Smith, B., CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. *Scientific Data*. vol 7 pp. 1-5, 2020.

[2] Donnelly, K., SNOMED CT: The advanced terminology and coding system for eHealth, *Studies in health technology and informatics*, vol 121 pp. 279-290, 2006.

Market Research using Big Data

Liam Fistos[1,2]
[1]M Science, New York, NY, USA
[2]Computer Science Department, Pace University, New York, NY, USA
lf32934n@pace.edu

Data Driven research and analytics is revolutionizing how companies, governments, and investors make decisions. Structured and unstructured data is parsed and aggregated in order to create actionable insights in near real-time. Metrics generated can be high level economic trends, all the way down to very granular customer cohorting. Although much of the work is still manual and/or human-guided, machine learning can be involved in a number of steps in the process.

- Anomaly Detection
- QA/QC
- Automated Parsing
- Metric Generation

This type of research will become more common as its value is proven and disruptive events force companies to rely on points of truth rather than past experience. Only a fraction of the potential data types and methodologies have been explored.

Classifying Knee Vibroarthrographic Signals Using Convolutional Neural Networks

Carmine Guida and Juan Shan
Computer Science Department, Pace University, New York, NY, USA
{cguida,jshan}@pace.com

Some methods for detecting Osteoarthritis may require invasive procedures. Vibroarthrography is a non-invasive technique which uses vibrations detected by accelerometers placed on the knee. This study converts a dataset of 1D signals into a 2D image and then implements a CNN model to classify the data. The dataset [1] used for this study contains 89 total cases (51 normal and 38 abnormal). For each case, the signal data was normalized and then converted into a monochrome 2D plot. The dataset is split into a balanced validation set (8 cases) and test set (14 cases). The remaining data is placed in the training set. The CNN model used can be seen in Figure 1. This model includes dropout layers along with early stopping based on validation loss in order to avoid overfitting. Additionally, 5-fold cross validation was used and the final F1 score was calculated as an average over 5 iterations of the experiment with different random initialization. An average F1 measure of 0.62 was achieved. While other works [2] utilizing the same dataset were able to achieve higher F1 scores with various machine learning techniques, the author plans to further work in this area by exploring possible preprocessing and adjustments to the model.

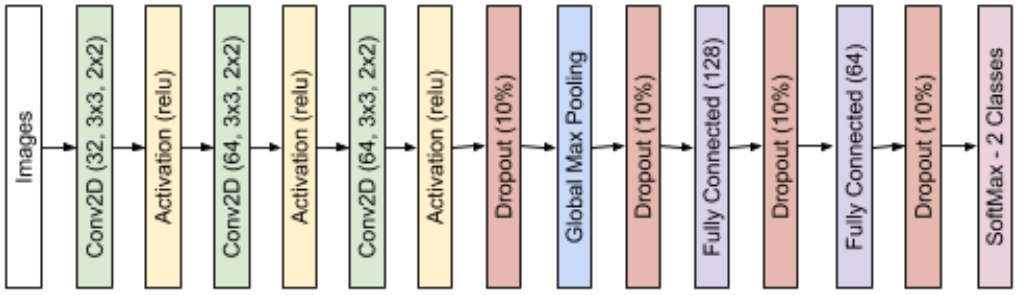


Figure 1: 2D CNN Model.

References

[1] Krishnan, S., Rangayyan, R. M., Bell, G. D., and Frank, C. B., "Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology," *IEEE Transactions on Biomedical Engineering*, 47(6):773-783, June 2000.

[2] Kraft, D., and Bieber, G., "Vibroarthrography using convolutional neural networks". In Proceedings of *the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2020 June, pp. 1-6.

Increasing Accuracy of Self-Driving Car and Avoiding Overfitting using Drop Out

Arun Kumar Yarram
Information Systems Department, Pace University, New York, NY, USA
ay28674n@pace.edu

The goal of these project is to increase the accuracy of Self-Driving car , so basically we have used Nvidia Architecture and made some changes to it as per our model, so the Autonomous cars takes input from camera, sensors and other features and store it in central computer and process output like steering angle the wheel should turn , How much acceleration it need to apply, when to turn on indicators, wipers as per requirement. In our case I am using the Dataset provided by SULLYCHEN, 40000 images as training set and we will be predicting only one output which is the at what angle the steering should turn. So here will be dealing with regression model. I will be using the architecture provided by Nvidia with slight changes like adding Drop-out with FCN so that we can avoid overfitting and increase accuracy.

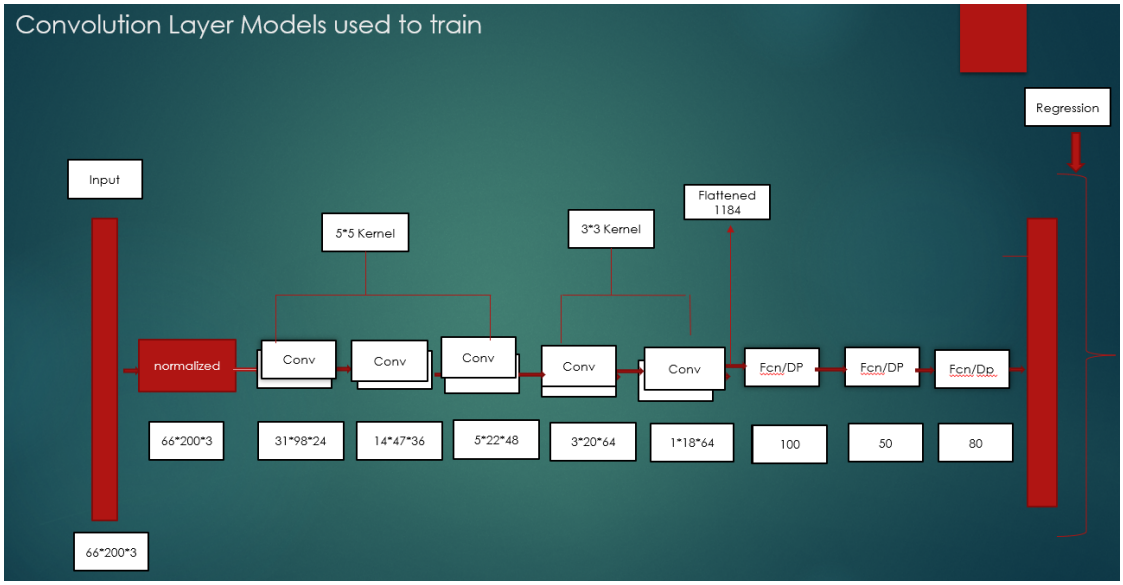


Figure 1: Architecture. The image is Taken from the source [**?**].

References

[1] Sully Chen, S., Dataset: https://github.com/SullyChen/driving-datasets

[2] End to End Learning for Self-Driving Cars Bojarski, M., Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J., and Zieba, K., NVIDIA. 2016

Automatic Detection and Classification of Leukemia using Microscopic Blood Images

Hetali Chavda and Vikas Sanhotra
Computer Science Department, Pace University, New York, NY, USA
{hc18137n,vs97960n}@pace.edu

Leukemia is a type of cancer that develops in the bone marrow. In leukemia patients, the bone marrow produces abnormal white blood cells [1]. An automated system for blood cell segmentation and classification is proposed in this study. The system analyses the microscopic images and remediates the drawbacks of manual checking. The dataset [2, 3] in this study contains a total of 134 Microscopic Blood Images of patients (75 with Leukemia and 59 without Leukemia). For detection and segmentation, the system extracts required parts of the images and applies the K-means clustering segmentation technique for the detection of white blood cells. Some of the morphological operations are used for removing noise and region filling. The Convolutional Neural Network (CNN) is applied for the classification. The system proposed for this study (as shown in Figure 1), is based on microscopic blood images for predicting if a patient has Leukemia or not. The system proposed is tested on an image dataset with 50.31% accuracy. The system has an overfitting problem, which can be improved in the future, by collecting more datasets and tune the model structure or fuse multiple models to improve accuracy.

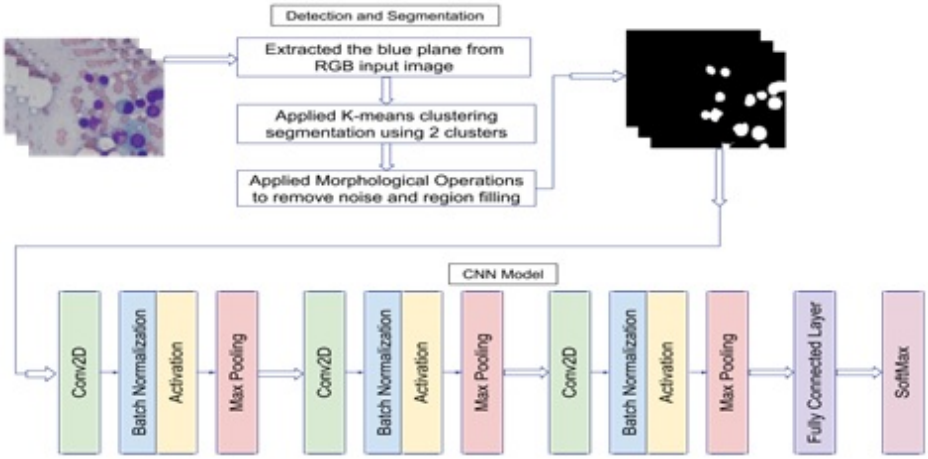


Figure 1: Proposed system

References

[1] Patel, N. and Mishra, A., Automated Leukaemia Detection Using Microscopic Images, *Procedia Computer Science*, vol 58, pp 635 - 642, 2015

[2] [4] Gupta, A. and Gupta, R.,. SN-AM Dataset: White Blood Cancer Dataset of B-ALL and MM for Stain Normalization. The Cancer Imaging Archive, 2019 https://doi.org/10.7937/tcia.2019.of2w8lxr

[3] Labati, R. D., Piuri, V., and Scotti, F., ALL-IDB: the acute lymphoblastic leukemia image database for image processing, in Proceedings of the *IEEE Int. Conf. on Image Processing* (ICIP), Brussels, Belgium, pp. 2045-2048, 2011

Extraction of Action Sequence Images From Blurred Shots

Anchal Singh and Harshada Chaudhari
Computer Science Department, Pace University, New York, NY, USA
{anchalsanjay.singh, harshadaanil.chaudhari}@pace.edu

The idea of the project previously added was, to develop an algorithm that combines multiple images generated from burst shots of an action of an object into a single image that clearly shows the full action [1]. The goal is to input a series of images in the form of burst images or video, into the system and output a single image representing the action which has a seamless background, multiple moments of the main subject and minimal noise around the subject [2].

While taking the series of images to create a single action image, we faced two challenges with different inputs. The series of images can be a burst shot of the images or the blurry images which could be shot by shaky hands. We proposed iterative approach to resolve the burst shots of the images problem. During this approach, we are using "background extraction" technique to extract one foreground image from the sequence images of burst shots at a time and that image will be processed through our proposed algorithm "Deghosting technique" [2] similarly each multiple desired output image will be taken from sequence of action images will then be processed and attached as a single image/panoramic view as an output. However, after applying the deghosting technique, the high quality blurry or shaky image problem can then be resolved using the "Multiple image deblurring" technique [3]. This algorithm removes all the noise from the images which were generated by shooting the video or clicked multiple images by shaky hands. In this technique each image will be processed separately to remove the blur effect from the images and removed portion of the image will be refilled by the background pixels/colors. In this paper, we are implementing the algorithm to extract the background and foreground images from the given burst images and merging all the action sequence of the object into one image. In future, we will be focusing on implementation of the solution to tackle the challenges with blur images or images taken by moving camera.

References

[1] Chen, S., Stabler, B., and Stantley, A., "Automatic Generation of Action Sequence Images from Burst Shots. Stanford University (2013)

[2] Szeliski, R. and Shum, H.. "Creating full view panoramic image mosaics and environment maps. in Proceedings of *the 24th annual conference on Computer graphics and interactive technique*, pp 251-258, 1997

[3] Ito, A., Sankarnarayanan, A. C., Veeraraghavan, A., and Baranuik, R. G., "BlurBurst: Removing Blur Due to Camera Shake using Multiple Images, Rice University

Image caption

Sichao Liu
Computer Science Department, Pace University, New York, NY, USA
sl39575n@pace.edu

Image Caption generation, is a composite of pattern recognition and machine learning technology that translating an Image into a Caption. The model commonly used for this task is a combination of two independent architectures, CNN (convolutional neural network) and RNN (recursive neural network). In this case, the LSTM (long & Short-term memory network) is a special RNN type unit that contains storage units for long-term memory and retention of information. The basic principle of the hybrid model: CNN is used to extract feature vectors of objects and their spatial information from the image, and then the feature vectors are input into the RNN network through a fully connected linear layer to generate sequence data (such as word sequence), and finally generate the description of the image.

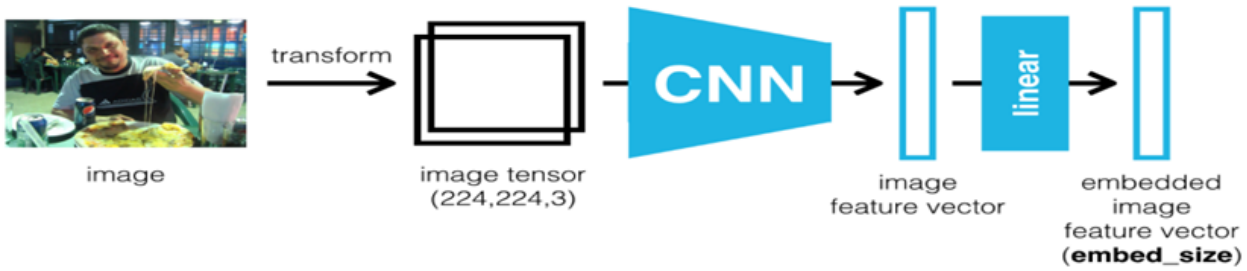


Figure 1: CNN Encoder Adopted from "Exploring LSTM"

References

[1] Karpathy, A., Johnson, J., and Fei-Fei, L., Visualizing and Understanding Recurrent Networks in Proceedings of *International Conference on Learning Representations*, San Juan, Puerto Rico, May 2016

[2] Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zeme, R., and Bengio Y., "Show, Attend and Tell: Neural Image Caption Generation with Visual Attention," in Proceedings of the *32nd International Conference on Machine Learning*, Lille, France, pp 2048–2057, 2015

Predicting Critical Region in Scene for Autonomous Vehicles

Vinay Babanrao More and Pratik Chaudhari
Computer Science Department, Pace University, New York, NY, USA
{vm57302n,pc42782p}@pace.edu

In autonomous vehicles analysing a complete camera view is computationally heavy task. In this work I aim to build a Realtime prediction model to predict which part of the scene is critical to focus. The goal is to build a computationally efficient model which works on only camera sensor to predict the critical situation for an autonomous vehicle. For this experiment building the own dataset is complex because critical driving moments are rare, gathering information for such circumstances is troublesome with the conventional data collection method like tracking eye movements of driver while driving. Humans get distracted easily, so it is important to focus on where the ideal driver should look at and not where the driver is looking at in a particular scene. To accomplish this goal, I am considering the methodology of ideal drivers focus while driving using Berkeley DeepDrive Attention (BDD-A) dataset. This dataset is based upon braking event videos which are selected from a large-scale crowd-sourcing driving video dataset. I built an attention prediction model that attends to pedestrians crossing the road and not on the pedestrians walking on the sidewalk.

Figure 1: Critical Regions in Scene.

References

[1] Alletto, S., Palazzi, A., Solera, F., Calderara, S., and Cucchiara, R., DR(eye)VE: A Dataset for Attention-Based Tasks with Applications to Autonomous and Assisted Driving, in Proceedings of *IEEE Computer Vision and Pattern Recognition Workshops* (CVPRW), pp 54-60, 2016

[2] Palazzi, A., Solera, F., Calderara, S., Alletto, S., and Cucchiara, R., Learning where to attend like a human driver, *IEEE Intelligent Vehicles Symposium* IV, 2017, pp 920-925

[3] Fernandez, P. D. M., and Peña, F. A. G., Ren, T. I., and Cunha, A., FERAtt: Facial Expression Recognition With Attention Net, in Proceedings of *IEEE/CVF Computer Vision and Pattern Recognition Workshops* (CVPRW), 2019, pp 837-846

[4] Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., and Durand, F., What Do Different Evaluation Metrics Tell Us About Saliency Models?, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 41, no. 3, pp. 740-757, 2019

Identification of the Number of Modes of a Data Distribution Based on the Kernel Density Estimation

Endri Zylali and Shashank Valmik Jadhav
Computer Science Department, Pace University, New York, NY, USA
{ez82613n,sj84593n}@pace.edu

This report evaluates different data distributions while processing a large number of datasets which can potentially have unimodal or multimodal distribution. According to dataset distribution in case of unimodal distribution which model can best approximates the dataset among uniform distribution, T-distribution, chi-square distribution, cauchy distribution, etc [1]. In case of multimodal distribution, automatic identification of number of modes and more granular descriptive statistic [2] is the main part of study in this report. This report provides estimation of probability density function by evaluating Histograms vs probability density function approximation, Kernel density estimations, Choice of optimal bandwidth: Silverman/ Scott/ Grid Search Cross Validation, statistical tests for unimodal distributions, DIP test for unimodality and Identification of the number of modes of a data distribution based on the kernel density estimation. The main purpose of this study is identification of the number of modes of a data distribution based on the kernel density estimation. While processing various large number of datasets which may potentially have different dataset distributions, we have confronted a few considerations which are:

- Is the data distribution unimodal and if it is the case, which model best approximates it (uniform distribution, T-distribution, chi-square distribution, cauchy distribution, etc.)?
- If the data distribution is multimodal, can we automatically identify the number of modes and provide more granular descriptive statistics?
- How can we estimate the probability density function of a new dataset?

Identifing the number of modesof a data distribution can be achieved by evaluating Histograms vs probability density function approximation, Kernel density estimations, choice of optimal bandwidth among: Silverman/ Scott/ Grid search cross validation, statistical tests for unimodal distributions, DIP test for unimodality and Identification of the number of modes of a data distribution based on the kernel density estimation.

References

[1] Xin Yi, Ekta Walia, Paul Babyn, *Generative adversarial network in medical imaging: A review* Department of Medical Imaging, University of Saskatchewan, Canada, 2019.

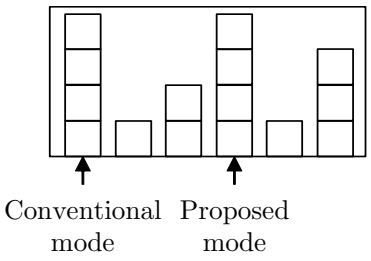
[2] Xin Yi, Ekta Walia, Paul Babyn, *Generative adversarial network in medical imaging: A review* Department of Medical Imaging, University of Saskatchewan, Canada, 2019.

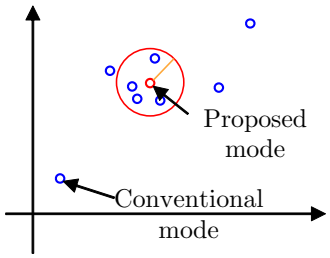
On the Density based Mode Within Radius

ParamPuneet Kaur Thind[1], Vaibhav Kumar Katturu[1], Teryn Cha[2], and Sung-Hyuk Cha[1]
[1]Computer Science Department, Pace University, New York, NY, USA
[2]Computer Science, Essex County College, Newark, NJ, 07102, USA
{pr54854n,vk38221n}@pace.edu, yan@essex.edu, scha@pace.edu

Statistical parameters such as center points are widely used to represent patterns in machine learning, e.g., EEG biometrics [1]. There are three popular center points: mean, median, and mode. Mean and median are pervasively used as features to represent patterns, but mode is not because of its weakness. One of the major problems in the conventional definition of mode occurs when the distribution is multimodal with tied multiple maximum points. Hence, we propose a modified version of mode, which utilizes the density within a radius.

$$\text{mode}(X) = \underset{x \in X}{\text{argmax}} \sum_{y \in X} N(x, y, r) \quad \text{where } N(x, y, r) \begin{cases} 1 & \text{if } d(x, y) \leq r \\ 0 & \text{otherwise} \end{cases} \tag{1}$$





The effectiveness of proposed mode is demonstrated in the emotion recognition application using the Electroencephalogram (EEG).

References

[1] Li, S., Cha, S.-H., and Tappert, C. C., "Biometric Distinctiveness of Brain Signals Based on EEG," in Proceedings of the *IEEE 9th International Conference on Biometrics Theory, Applications and Systems* (BTAS), p 1-6, 2018

Food Detection Using Statistics of Pairwise Local Features

Sanath Gholap, Tushar Rakholiya and Ramesh Kyasaram
Computer Science Department, Pace University, New York, NY, USA
{sanathsadanand.gholap, tushar.m.rakholiya, rameshtamanna.kyasaram}@pace.edu

It is difficult to identify food because food products are reconfigurable items that have significant variations in appearances. Automatic recognition of food is evolving as a significant thing in object recognition because of demands for improved nutritional assessment to overcome obesity. This will help people to have a better understanding of the nutritional content of their food choices and provide medical practitioners for objective measures for food intake of their patients. When disclosing the food that they eat, people are not very reliable. We explore methods for automatically identifying foods based on their presence, as an alternative to manual recording. Unfortunately, based on aggregating statistics of descriptive local characteristics, the traditional objects recognition approach severely fails since food products are deformable and shows substantial differences in presentation in intra-class.

We know that analyzing the spatial associations between different ingredients is the secret to recognizing food. We propose a new representation of food products that calculates pairwise statistics between local characteristics determined by extracting features into eight types of ingredients over a soft pixel level. We accumulate these statistics in a multi-dimensional histogram, which is then used as a function vector for a distinguishing classifier. Our test demonstrates that in recognizing food, the proposed representation is substantially more reliable than the existing methods.

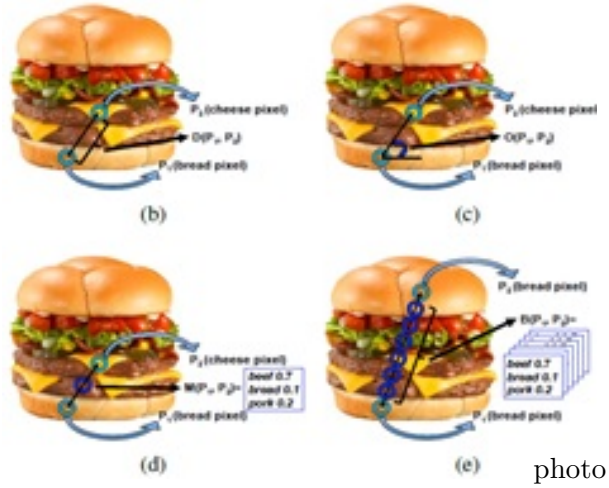


photo from [1]

Figure 1: Exploiting spatial relationships between ingredients using pairwise feature statistics.

References

[1] Yang, S., Chen, M., Pomerleau, D., and Sukthankar, R., Food recognition using statistics of pairwise local features, in Proceedings of IEEE *Computer Vision and Pattern Recognition*, 2010, pp 2249-2256

[2] Chen, M., Dhingra, K., Wu, W., Yang, L., Sukthankar, R., and Yang, J., PFID: Pittsburgh fast-food image dataset, in Proceedings of 16th IEEE *International Conference on Image Processing* (ICIP), 2009, pp 289-292

Improving CNNs with the help of Data Augmentation and Boosting

Soham Save
Computer Science Department, Pace University, New York, NY, USA
soham.k.save@pace.edu

A Convolutional Neural Network (CNN) is a deep learning algorithm which is used to analyze images. A CNN takes images as an input in form of arrays, assigns weights and biases. For learning, to the parts of the image so that it can differentiate one part from the other. The main reason for using CNN is that it can capture spatial and temporal dependencies of an image with the help of filters [1]. Image Augmentation is a commonly used technique when an image dataset size for CNN is less. Image Augmentation is used to increase the dataset by creating modified versions of the images present in the dataset. This in turn will improve the performance and the ability of CNN to generalize. Image Augmentation has been really proved to be helpful in CNN performance [2].

However, we can boost the performance of CNN, so as the name suggests we use boosting algorithms. Boosting algorithms convert weak learners into strong learners. Boosting out performs simple CNN. In this experiment, I boosted my CNN by extracting features from the mid layer, a Dense layer just before the softmax layer and feeding them to XGBoost, one of the Boosting algorithms. After providing Boosting, the results improve vastly [3].

References

[1] Valueva, M. V., Nagornov, N. N., Lyakhov, P., Valuev, G. V., and Chervyakov, N. I., Application of the residue number system to reduce hardware costs of the convolutional neural network implementation, Mathematics and Computers in Simulation, vol 177, pp. 232-243 2020.

[2] Shorten, C. and Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, SpringerOpen vol 6, num 1, pp 1-48, 2019.

[3] Zhou, Z. H., Ensemble Methods: Foundations and Algorithms, Chapman & Hall/CRC data mining and knowledge discovery series, CRC Press, 2012

A Weighted Preferential Vote for Pattern Classification

Sayema Islam and Sung-Hyuk Cha
Computer Science Department, Pace University, Pleasantville, NY, USA
{si54044p, scha}@pace.edu

Various methods of pattern classification, popularly weighted and preferential voting, rely on ensembles of learners, which consistently outperform simpler, single classifiers in terms of accuracy [1, 2]. However, when one considers the existence of weight in ensembles, that some individual classifiers within a group may outperform others, this raises a question of whether or not it is possible to compare the relative performances of CNNs within ensembles to one another. And while this determination of a networks certainty has been examined by Jimenez and Walsh in their study with dynamically weighted voting, the study of this same performance measure in the realm of preferential or ranked voting remains largely unexplored [2].

Thus, by examining various methods of preferential voting and choosing one most suitable to be translated into a system of weighted preference, this study puts forth a proposal to apply the logic of a chosen preferential voting algorithm to create weights for various classifiers based on their relative accuracies. In doing so, a new importance of preference metric is put forth, which then paves the way for a new method of voting that can be applied to ensembles of CNNs to compare their relative performances in ranked classifications, a consideration that proves especially important in illness and disease classification.

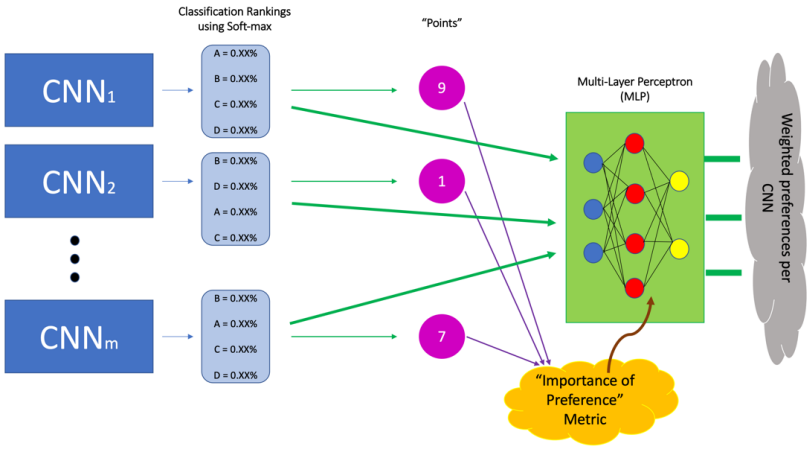


Figure 1: Proposed two-stage ensemble for testing viability of new measure.

References

[1] Kattaru, V. K., Thind, P. KT. Cha, T., Cha, S.-H., " A New Ensemble Method for Convolutional Neural Networks and its Application to Plant Disease Detection, In Proceedings of the *16th International Conference on Machine Learning and Data Mining*, (MLDM), Amsterdam, The Netherlands, July 20-21, 2020, pp. 187-196

[2] Jimenez, D., Dynamically weighted ensemble neural networks for classification, in Proceedings of *the IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence* (Cat. No.98CH36227), 1998, pp 753-756

Improving CNNs with the help of Data Augmentation and Boosting

Siddharth Urvesh Jhaveri
Computer Science Department, Pace University, New York, NY, USA
sj17917n@pace.edu

A Convolutional Neural Network is a machine (deep) learning algorithm that is used to work with complex data, depicting the logic of a human brain to some extent and try to infer predictions using Convolution, Pooling and Densification to draw classifications. It is most widely used on images, due to its convolution + RelU architecture. The architecture of the entire network will define the accuracy of the model built.

Image Augmentation is a technique used to artificially increase the dataset and its variety by using the current images [1] and changing them by using certain parameters, such as rotation, skew, horizontal/vertical flip, etc.

The purpose of this project is to inculcate this augmentation as data feed into a developed neural network and feed it as data for the next epoch of the network. This will help boost the performance of the model being built and keep a certain amount of ambiguity in the data being fed every epoch. The dataset being used is available to all [2] and consists images of sceneries and the goal is to identify which 1 of 6 it truly is. The goal will be to observe if image augmentation can serve as a viable boost for building a neural network or not.

References

[1] L. Taylor, L. and Nitschke, G., Improving Deep Learning using Generic Data Augmentation in Proceedings of the *IEEE Symposium Series on Computational Intelligence* (SSCI), 2018, pp 1542-1547

[2] Bansal P., Intel Image Classification: Image Scene Classification of Multiclass, Dataset: https://www.kaggle.com/puneet6060/intel-image-classification, 2019

Friend Recommendation System using Ensemble Voting and Distance Heuristics

Viswanathan Appakkudal Ramani, Dhariya Harsh Parikh, and Yashwanth Varre
[1]Computer Science Department, Pace University, New York, NY, USA
[2]Computer Science, Essex County College, Newark, NJ, 07102, USA
{va99912n,dp08607n,yv28667p}@pace.edu

Friend Recommendation System helps to forecast future friends by analyzing social network using an ensemble of classifiers and weighted voting strategy. Social network analysis is a mechanism through which graphs are analyzed to find patterns and deduce relationships between multiple nodes. Three-person P1, P2, P3 form a length-3 following chain if a person P1 follows person P2 and person P2 follows P3. Given a length-3 following chain, it may make sense to recommend to P1 to follow person P3. It may make even more sense doing that if there are several 3-following chains between P1 and P3.

We considered the Facebook dataset provided by Stanford which contains 4039 nodes and 88234 edges and makes a directed network. We cleaned the graph by eliminating isolated nodes, and nodes with fake edges. We extracted features such as Jaccard Similarity, Adamic-Adar Index, Preferential attachment, Resource Allocation Index, and Common Neighbors [1]. Common neighbors find the number of mutual friends between two persons. Jaccard similarity compares members for two sets to see which members are shared and which are distinct. Resource Allocation Index is a measure that calculates the amount of resource that a node can send to the destination node via its neighbors. Preferential attachment measures the influence of a node over the other. Adamic-Adar index measures the number of shared links between two nodes. We modeled the features using an ensemble of ANN, Logistic regression, and Nave Bayes classifiers using a weighted voting mechanism [2]. Classifiers are weighted based on their accuracy score; the classifier with maximum accuracy gets the highest weights, and weights reduce for a classifier as the accuracy reduces. And, if the accuracy is less than a threshold, or if the classifier overfits/underfits the data, the classifier is eliminated, thereby keeping only significant classifiers.

References

[1] Fire, M., Tenenboim, L., Lesser, O., Puzis, R., Rokach, L., and Elovici, Y., Link Prediction in Social Networks Using Computationally Efficient Topological Features, in Proceedings of *IEEE Third International Conference on Privacy, Security, Risk and Trust* and *IEEE Third International Conference on Social Computing*, 2011, pp 73-80

[2] Chebrolu, S., Abraham A., and Thomas, J. P., Feature deduction and ensemble design of intrusion detection system, *International journal of computers and security*, vol. 24, no. 4, pp. 295-307, 2005

Audio Visual Emotion Recognition

Pratik Chaudhari, Vinay More, and Pradnya Sangoram
Computer Science Department, Pace University, New York, NY, USA
{ps29394p, pc42782p, vm57302n}@pace.edu

Understanding what an individual is encountering from her edge of reference is basic in our regular daily existence. Therefore, one can imagine that machines with this sort of capacity would associate better with individuals. In any case, there are no current frameworks fit for comprehension in detail individuals' enthusiastic states. In this paper we try to build a method capable of recognizing feelings in the same way as humans do and has a lot of potential applications in Human-Computer Interaction (HCI), human-assistive technologies and online education, among others. In this paper we use the "Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)", a dataset of images containing 24 professional actors (12 female, 12 male), expressing two lexically coordinated explanations in a neural North American articulation. With the RAVDESS dataset, we jointly trained a Convolutional Neural Network (CNN) model to analyze the actor and the entirety scene to perceive rich information about emotional states. Through this paper, we show the significance of thinking about the context for perceiving individuals feelings in images. We started our research to build a fusion network in which we will implement a new voting method for predicting the emotion of video based on the video and the speech. Video is a stack of images and Convolutional Neural Network performs well in image classification, as well as Speech is a sequential data and we are using Recurrent Neural Network to get the probabilities of emotion from speech. In initial phase we are focussing on the Visual data and performing several experiments to find the important feature that can play vital role in recognising emotion from video frames. Initially we considered this as an image classification problem but upon performing few experiments it turned out not to be a simple image classification problem. The important feature in facial emotion recognition is facial landmarks of the person and we are considering extracting facial landmarks and using it as input for our Convolutional Neural Network.

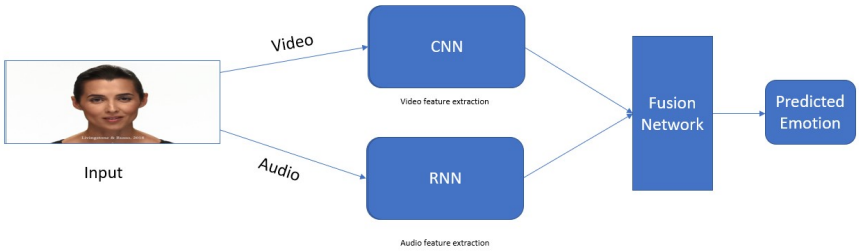


Figure 1: How is this person feeling?

References

[1] Livingstone S. R. and Russo F. A., The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. PLoS ONE 13(5): 2018

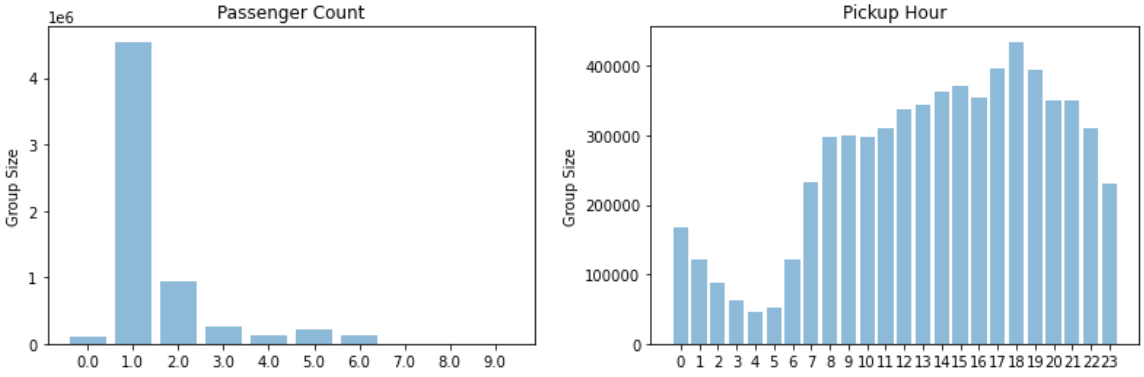
[2] Fernandez, P. D. M., Peña, F. A. G., Ren, T. I., and Cunha, A., FERAtt: Facial Expression Recognition With Attention Net, IEEE Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp 837-846

[3] Chollet, F., Xception: Deep Learning with Depthwise Separable Convolutions, IEEE Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 1800-1807

Deep Analysis of NYC/Limo System

Xu Zhang and Lang Gong, and Ou Ge
Computer Science Department, Pace University, New York, NY, USA
{xz48433n, lg16827n, go78678n}@pace.edu

Yellow cab is an icon of New York city with a history of more than 70 years. Together with Green Taxi, FHV and High Volume FHV, it creates a robust and versatile system to support the transportation system for NYC and its five boroughs from time to time. The New York City Taxi and Limousine Commission (TLC), created in 1971, is the agency responsible for licensing and regulating New York City's medallion (yellow) taxis, street hail livery (green) taxis, for-hire vehicles (FHVs), commuter vans, and paratransit vehicles. The TLC collects trip record information for each taxi and for-hire vehicle trip completed by our licensed drivers and vehicles. For each month, the dataset consists four csv files for yellow taxi, green taxi, FHV and High volume FHV. For each file, it includes columns like pickup/dropoff time, pickup/dropoff locations, time duration, mileage, breakdown of fares and so on. In NYC, the traffic, especially during peak hours, is a big headache. As the major participants in NYC traffic, NYC Taxi might be able to give us some insights about NYCs traffic condition. Even further, we can create traffic predictor to predict the travel time given some parameters. Combined with the large data of yellow and green taxis in the New York City, the traffic data is processed based on the MATLAB and Spark, which is a distributed data processing platform, and use the K-means clustering algorithm to analyze the center point of boarding location. For the exploratory analysis we made different charts as follows:



References

[1] Venkata, A., "K Nearest Neighbors and implementation on Iris data set," May 18,2019.

[2] Vik, P., "21 places to find free dataset for data science project," Sep 16, 2020.

[3] Adbul, Q., "Linear regeression and exploratory data analysis using R on happiness database 2019," Dec 15, 2019.

Research of COVID-19 Impact and Trend in USA using Statistical Approach

Artsiom Nasledniakau
Computer Science Department, Pace University, New York, NY, USA
art.nasledniakau@gmail.com

Research tends to explore a public dataset "United States COVID-19 Cases and Deaths by State over Time" from Centers for Disease Control and Prevention to visualize impact and trend in the USA on population. The research used computer science methods such as aggregation, pattern recognition, feature classification, data regression, clustering and attribute selection.

A difference-by-day algorithm was developed and applied since raw data from the dataset had linear forms. A new parameter "recovery" was introduced to cross-validate trend results which was also used as the metric to compare results from linear and isotonic models. Both models showed similar results with a little deviation.

As part of the research, data from the dataset have been grouped by US state and each US state has been classified by "mean" feature. The classified result showed average and inflated groups. Also, each US state had been plotted to map with a metric density color. Besides, a k-means cluster was generated for US states.

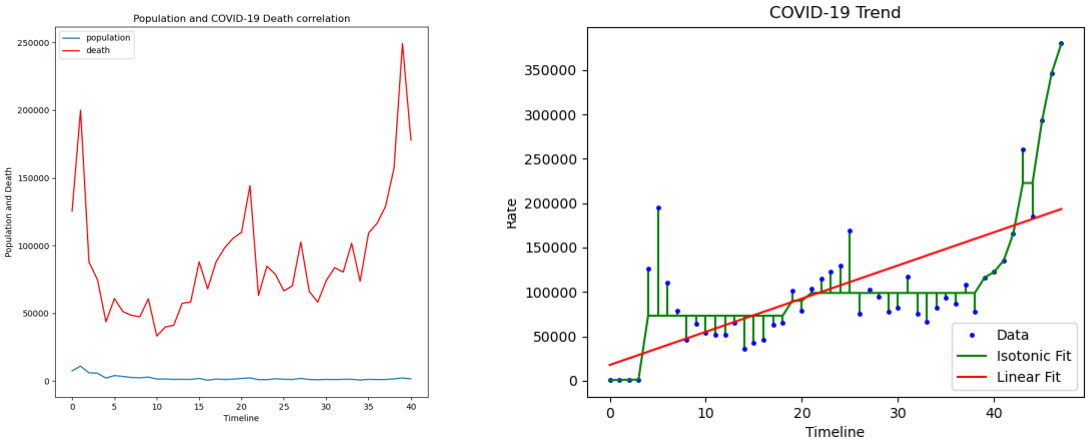


Figure 1: COVID-19 impact of USA population and COVID-19 case trend

References

[1] CDC - United States COVID-19 Cases and Deaths by State over Time, 2020 - https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36/.

[2] Census - U.S. and World Population Clock, 2020 - https://www.census.gov/popclock/

[3] Kumar, N. and Susan, S., COVID-19 Pandemic Prediction using Time Series Forecasting Models, in Proceedings of the *11th International Conference on Computing, Communication and Networking Technologies* (ICCCNT), 2020, p 1-7

Using Abstarct Syntax Tree with CodeBERT to Improve BLUE Score in Zero-Shot Code Documentation Generation

Dhruvil Gandhi
Computer Science Department, Pace University, New York, NY, USA
dgandhi@gmail.com

Multiple pre-trained models have been designed to work with Natural Language Programming Language pair bimodal data. In this paper, a pre-trained model is proposed using CodeBERT model with code documentation data and corresponding code as Abstract Syntax Trees. The model is developed with Transformer based neural architecture. Result are evaluated for zero-shot code documentation generation using clean open-source GitHub Repositories data provided by CodeBERT [1] and its performance is compared with BLEU score of state-of-the-art models. [1, 2, 3]

References

[1] Feng, Z., Guo, D., Tang, D., Duan, N., Feng, X., Gong, M., Shou, L., Qin, B., Liu, T., Jiang, D., and Zhou, M., CodeBERT: A Pre-Trained Model for Programming and Natural Languages, 2020

[2] Ott, M.,. Edunov, S, Baevski, A., Fan, A., Gross, S., Ng N., Grangier, D., and Auli, M., fairseq: A Fast, Extensible Toolkit for Sequence Modeling, in Proceedings of *NAACL-HLT 2019: Demonstrations*, 2019

[3] Alon, U., Levy, O., and YahavE., code2seq: Generating Sequences from Structured Representations of Code, CoRR, https://dblp.org/rec/journals/corr/abs-1808-01400.bib

Index of Participants

A - Authors CC - Conference Chairs PC - Program Committee LA - Local Arrangement