ISSN: 1947-3931 Volume 12, Number 3, March 2020

Scientific
Research
Publishing

Engineering

ISSN: 1947-3931

https://www.scirp.org/journal/eng

Journal Editorial Board

ISSN Print: 1947-3931 ISSN Online: 1947-394X

https://www.scirp.org/journal/eng

Editor-in-Chief

Editorial Board

Scientific Research Publishing

Engineering, 2020, 12, 151-227
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

Table of Contents

Volume 12 Number 3 March 2020

Engineering (ENG)
Journal Information

SUBSCRIPTIONS

The *Engineering* (Online at Scientific Research Publishing, https://www.scirp.org/) is published monthly by Scientific Research Publishing, Inc., USA.

Subscription rates:
Print: $89 per issue.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: eng@scirp.org

Scientific Research Publishing

Engineering, 2020, 12, 151-165
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

Creating a Dataset to Boost Civil Engineering Deep Learning Research and Application

Murad Al Qurishee[1], Weidong Wu[2], Babatunde Atolagbe[3], Joseph Owino[4], Ignatius Fomunung[5], Mbakisya Onyango[5]

[1]Graduate Transportation Associate, Tennessee Department of Transportation, Nashville, Tennessee, USA
[2]University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
[3]Transportation Engineer, Maryland Department of Transportation, Baltimore, Maryland, USA
[4]UC Foundation Professor and Department Head, Department of Civil & Chemical Engineering, The University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
[5]Department of Civil & Chemical Engineering, The University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
Email: mgp216@mocs.utc.edu, Weidong-Wu@utc.edu, zxc259@mocs.utc.edu, JosephOwino@utc.edu, Ignatius-Fomunung@utc.edu, MbakisyaOnyango@utc.edu

How to cite this paper: Al Qurishee, M., Wu, W., Atolagbe, B., Owino, J., Fomunung, I. and Onyango, M. (2020) Creating a Dataset to Boost Civil Engineering Deep Learning Research and Application. *Engineering*, **12**, 151-165.
https://doi.org/10.4236/eng.2020.123013

Received: January 20, 2020
Accepted: March 3, 2020
Published: March 6, 2020

Abstract

With cutting edge deep learning breakthrough, numerous innovations in many fields including civil engineering are stimulated. However, a fundamental issue that civil engineering research community currently facing is lack of a publicly available, free, quality-controlled and human-annotated large dataset that supports and drives civil engineering deep learning research and applications on such as intelligent transportation including connected vehicle, structural health monitoring, and bridge inspection. This paper is a general discussion about demanding needs and construction of a long-anticipated dataset for researchers and engineers in civil engineering and beyond for providing critical training, testing and benchmarking data. The establishment of such a free dataset will remove a major hurdle and boost deep learning research in civil engineering and we hope this work will urge researchers, engineers, government agencies and even computer scientists to work together to start building such datasets. A framework has been developed for the proposed database. Also, some pilot study databases were developed for concrete crack detection, pavement crack detection using normal and infrared thermography, as well as pedestrian and bicyclist detection. A convolution neural network model called Faster RCNN was deployed to check the detection accuracy and a 98% detection accuracy of the proposed datasets was obtained.

Keywords

Dataset, Deep Learning, AI, Civil Engineering, ITS, Transportation

1. Introduction

With the breakthrough of deep learning due to advances of hardware such as GPU and Google cloud TPU chip [1], available large datasets as ImageNet [2] and benchmarks and algorithm improvements such as better activation functions, better weight-initialization schemes and better optimization schemes [3], numerous innovations have been unveiled and researchers from various communities are being excited. Civil engineering researchers have applied this novel computer technology to damage detection of structure [4] [5], concrete crack [6] [7], bridge structural components [8], pavement [9], tunnel [10], transmission tower [11] and roof [12]. Some of the most significant commercial products built on deep learning are Apple HomePod®, Amazon Echo®, and Google Home®.

One big force that drives the advance in computer vision, machine learning and deep learning is publicly available large and high-quality datasets that stand behind various competitions hosted by Kaggle platform that challenges data science and Artificial Intelligence. One example of such competition is Image Large Scale Visual Recognition Challenge (ILSVRC) that is hosted annually in order to challenge computer algorithms for object localization/detection from images and videos [13].

Data has been playing an irreplaceable role behind recent explosions and boom in artificial intelligence (AI). The key element or foundation of deep learning application and research is the dataset that supports training and testing deep neural networks and learning skills. In computer vision and machine learning (ML) research community, there are openly available datasets like ImageNet [2] that supported the pioneering deep learning publication AlexNet [14] and spawned the current AI boom, MS COCO [15] for objection, segmentation and captioning, CIFAR-10 [16], a labelled subset of 80 million tiny images that support ML [17], and MNIST [18] for handwritten digits. It took tremendous efforts, time and investment to build such large-scale datasets. For example, MS COCO researchers utilized more than 70,000 worker hours [15] to label and annotate millions of images and object instances. The data is so important to AI and deep learning, that it was reported very recently in June 2018 that the President's most senior technology advisor claimed that the White House may consider releasing some government data to push AI research [19].

However, the open dataset is of limited value to civil engineering research community. Most of the research papers in civil engineering fields used their proprietary datasets which are generally quite small compared to the above general-purpose datasets. While the success of deep learning owes to availability of large-scale labeled data [20], it would be less advantageous if only small datasets were used. Sun *et al.* [20] concluded that the performance of vision tasks increases logarithmically if the size of dataset gets large. If research data could be widely shared, a larger dataset would be created with fewer efforts. But researchers are reluctant to share their data due to legal issues and many other barriers

[21].

Various techniques have been proposed to solve the issue of small datasets for training and testing deep neural network. The creator of Keras deep learning library Chollet [22] proposed a few simple techniques to address small dataset for image classification problem, which includes data augmentation, and transfer learning [23]. Nevertheless, transfer learning only works for deep learning when the model features extracted from a large dataset from pre-training are general. In this sense, use for example a pre-trained ConvNet on ImageNet for transfer learning of crack and damage of concrete is still questionable and investigation must be done to properly evaluate the feasibility of using this technique for civil engineering applications purpose.

In the rest of the paper, a reader will expect the discussion of the application of AI in Civil Engineering research and how important a good database makes the AI techniques more efficient for the users. In the following section, a dataset structure is proposed and discussed how it can improve the traditional database limitations. At the end, some proposed databases were developed and tested using convolution neural network (CNN) models.

2. Civil Engineering Research and Application Needs

Modern computer technologies are making civil engineering, one of the oldest engineering disciplines, smarter and more intelligent. Worldwide cities and governments are launching smart cities [24] initiative. Intelligent Infrastructure can be one of the key characteristics that supports smart city initiatives. Intelligent infrastructure addresses intelligent transportation [25] [26], smart buildings and structures [27] [28], smart bridges and tunnels [29], smart pavement monitoring system [30], and etc.

The USDOT's Intelligent Transportation Systems (ITS) ITS Strategies Plan 2015-2019 outlined the goal of "Realizing Connected Vehicle (CV) Implementation" and "Advancing Automation" as the primary technological drivers of ITS [31]. Connected vehicle applications build an interoperable wireless communication network through dedicated short-range communications (DSRC) [32] or possibly on the deploying 5G telecom network in the near future, which collects vehicles, infrastructure (as traffic lights), and wireless devices (as cell phones) to prevent vehicle crash, improve mobility by reducing delay and congestion and benefit environment by cutting emission. In September 2016, US DOT selected New York City, Tampa-Hillsborough and Wyoming to launch its Connected Vehicle Pilot Deployment Program, which serves an initial effort to deploy, and test the cutting-edge CV technology [33]. Connected vehicle research has still been its early stage, which relies and demands sizable datasets that can serve benchmarking data for developing critical algorithms. Valuable, publicly available, selected and well-organized data related will be always appreciated in order to help advance fundamental knowledge and provide strong and timely support to connected vehicle research.

Allocation of Federal transportation funds and transportation infrastructure management and planning requires traffic monitoring, vehicle count, and classification. Very affordable video devices and State-of-the-art deep neural network based computer techniques/algorithms such as YOLO [34], and Faster R-CNN [35] can provide simple, manageable, cost-effective, real-time solution to any type of traffic counting and classification problem. A large dataset with images of all types of motorized and nonmotorized vehicles, pedestrians and bicyclists will ensure a well-trained deep neural network for detecting classifying and counting traffic.

Structural health monitoring [36] to buildings, bridges, tunnels, and any other civil infrastructure will provide a real-time and preventative strategy to identify and monitor potential damage to a structure. Application of wireless sensor network and wireless smart sensors [37] is the recent trends and future of civil infrastructure health monitoring. Trained deep neural network can be a strong candidate to automatically process the collected ambient vibrations, wind, strain, displacement data for structural damage detection and condition health assessment. However, training a deep neural network calls for a reliable large dataset. Other potential application of deep neural network can be dam and nuclear power plant concrete structure health monitoring, which may safeguard welfare and lives of hundreds of thousands of citizens.

Another critical civil engineering need is bridge inspection automation. With more than 56,000 or 9.1% structurally deficient bridges in the US, more and more bridges may require even shorter inspection interval than the basic 2-year requirement, which means more inspection and maintenance efforts, and higher costs and more dangerous works. With affordable Unmanned Aerial Vehicle (UAV) and deep learning computer technology, it is a trend to partially or completely replace vision inspection. Many civil engineering researchers including the authors of this article and computer scientists are working enthusiastically towards automation of bridge inspections. Recently we proposed a framework of coupling UAV and deep learning for civil infrastructure condition assessment. One of the major challenges civil engineering community face in applying deep learning in bridge inspection is shortage of an image dataset that has good representation of all bridge components to be inspected [38].

Deep learning has also been attempted to solve time-series based real-world applications. Researchers recently applied this computer technology to predict and forecast traffic flow [39] [40] in order to study traffic congestion and delay, and estimate building energy consumption [41] [42] to eventually develop smart grids. The traffic data used in these studies for training and testing are obtained from open data portal provided by State DOTs. And the electric power consumption data was collected from an individual residential customer. Not many publications have been found in these research areas partly due to insufficient quality data.

In summary, deep learning can transform many aspects of conventional civil engineering: structural health monitoring of important infrastructure in damage

detection automation, bridge inspection, intelligent transportation and connected vehicle, road condition assessment, traffic counting and planning, traffic flow prediction and building energy consumption estimation. We anticipate more and more exciting and meaningful applications of deep learning in civil engineering with large dataset and advanced deep neural network architectures available soon.

3. The Proposed Dataset

This section presents the difference of the proposed dataset from the existing ones, what should be included in the proposed dataset, how it can possibly be organized, the data collection methods and finally tools to build such a dataset.

3.1. Difference from Existing Datasets

The image datasets ImageNet [2] built by researchers from Princeton University and MS COCO [15] created by Microsoft provide good examples on how such datasets can be constructed, organized and made available to the civil engineering community and beyond. We can generally follow in the footprints of ImageNet and MS COCO. Nevertheless, ImageNet currently offers only images while the proposed dataset expects to include time-series data such as traffic flow, building energy consumption, connected vehicles data as well. The time-series data such as connected vehicle data [43] including basic safety message information as position, motion, vehicle size, road coefficient of friction, lights status and so on can be huge, which makes collect and store data very expensive. Furthermore, dataset as ImageNet hosts mostly daily life objects other than discipline specific targets. The needs for civil engineering applications are far beyond recognition and classification of daily life objects. Workers to label such images may need to receive Safety Inspection of In-Service Bridges training provided by FHWA National Highway Institute in order to produce quality labels for images to be used for bridge inspection. In addition, data in the proposed dataset rely on not only internet and contribution from researchers, but also government agencies. Even though government open data site data.gov offers tremendous amount of data, finding the right data for state-of-the-art in civil engineering deep learning research is without much luck.

3.2. Data Structure

The popular ImageNet was created with a hierarchical structure according to WordNet [44] that is a large lexical database of English words grouped into synonym set or synsets. Compared to ImageNet, the proposed dataset is more discipline (civil engineering) specific, it would be more appropriate to organize the dataset referring to government published documents, national standards or widely accepted classifications in the discipline. As seen in **Figure 1** that depicts the hierarchical structure of the proposed dataset, subtrees of class bridge are organized following National Bridge Inspection Standards and Bridge Inspector's

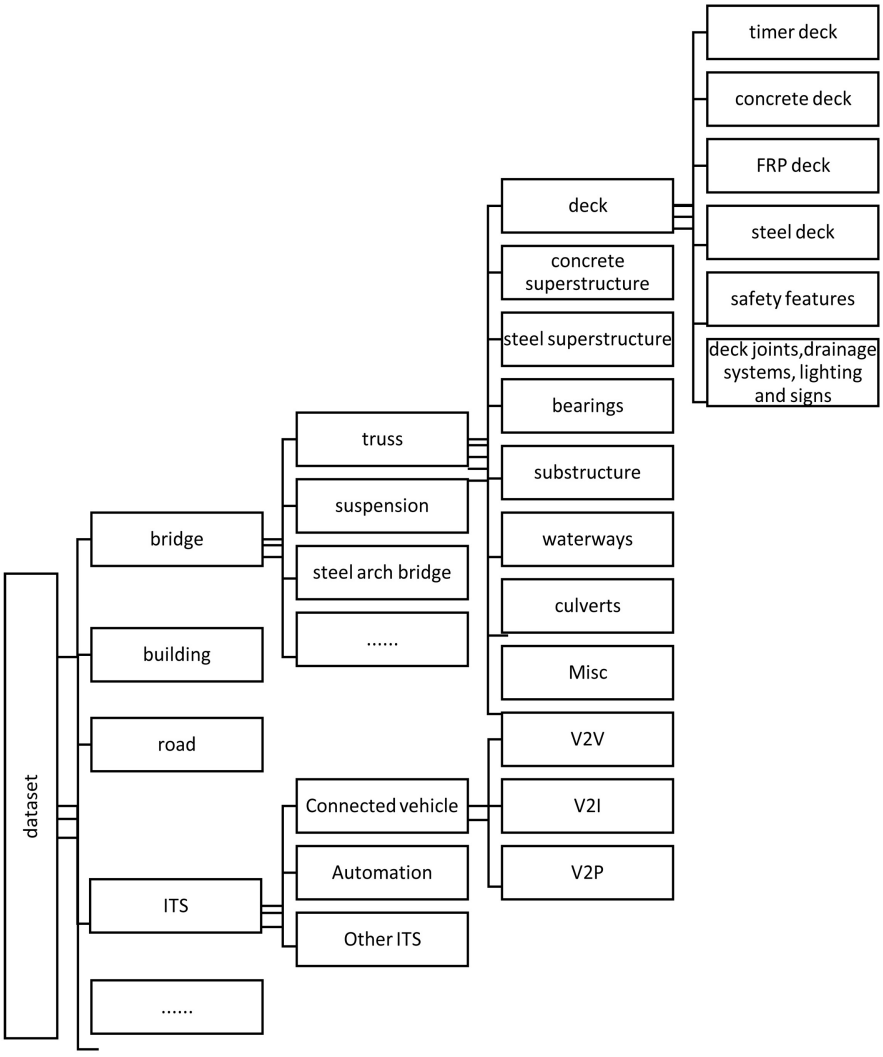


Figure 1. The hierarchical structure of the proposed dataset.

Reference Manual [45].

3.3. Data Collection

Data should have good representation and cover a broad range of research and application to serve the cutting-edge deep learning research in civil engineering. Labelling of images can be a daunting task, which may be completed in multiple ways such as web-based annotation LabelMe [46], Amazon Mechanical Turk (AMT) and even with a computer game [47].

Humans who receive special trainings may need to label some types of data such as images for highway bridge inspection.

The following is a list of various methods that may be used to collect data for the proposed dataset.

- Data mining online resources

Internet based data collection [48] is a relatively easy and cheap way to collect a large thus more representative data. Both ImageNet and MS COCO collect

images from internet. It is not a challenging task to use python script to automatically extract and scrape data. However, data collected from internet may be satisfying for daily life applications such as detection and recognition of cats and dogs, they may not always meet the demands of specific civil engineering research and applications. Another concern of using internet-based data is piracy and copyright protection. Fortunately, non-commercial use of data for education, and research is generally allowed. For example, researchers and educators may download images acquired from web by ImageNet for non-commercial and/or educational purchases under certain conditions and terms.

- Image data may be obtained by querying several image search engines such as Google Images, Bing Images, and Flickr. For example, MS COCO collected non-iconic images from Flickr. Kaggle can also be a good source to identify good data.
- Request data from DOT and local state DOTs and other government agencies.

This could be the best approach to obtain high quality data. The Freedom of Information Act (FOIA) [49] is a Federal law that gives individuals the right to access to any US federal agency records unless the agencies the release is prohibited by law or protected by nine exemptions, which means we may not necessarily be able to request all the data of interest.

There are tremendous amount of data and we first need to identify and decide what are the most valuable data to be requested that can be potentially used for deep learning studies. One example of valuable data can be images taken by bridge inspectors owned by state DOTs. These images may serve high quality data for supervised training and testing deep neural network for bridge inspection purpose.

- Collect and archive as traffic and data from publicly available open data portal provided by State DOTs and other agencies. **Table 1** shows a few examples of publicly available free traffic data portal provided by state DOTs. The proposed dataset website should extract data from those data portals, organize and classify for deep learning traffic flow research use, which may need heavy involvement of data cleansing to better serve the deep learning research needs.
- Promote and encourage share of data.

High quality data is the heart of any research work and excellent data builds the best possible foundation for deep neural network related publications. However, individual author may have their small datasets and look for even larger dataset for their use. "Take one, return one" (a researcher may download data if they contribute) may encourage share of data among peer researchers.

- Launch competitions based on the proposed dataset to help advance and develop better algorithms for civil engineering deep learning research and application.

Kaggle sets an example of providing predictive modeling competition platform to solve a wide variety of problems in different fields of computer science,

Table 1. Selected traffic data portals provided by state DOTs [7] [38] [50]-[63].

Data Portal	States and Agencies	Comments
PeMS [64]	CA	Real-time data by 40,000 detectors across CA. Archived historical data with a wide variety of information available; free
NYDOT Real Time Traffic Information	NY City	Real-time traffic cameras; free
Smart Way [65]	TN DOT	four transportation management centers, 517 cameras, 174 message signs, 1015 roadway detection systems and 49 video detection systems in the four largest cities; real-time traffic cameras
TADA	GA DOT	Dynamic mapping interface and other data formats to present data collected
ITS Public Data Hub [66]	US DOT	ITS data

computer vision, medicine etc. With emerging technologies such as connected vehicle, numerous challenges have been remaining and competitions based on the proposed datasets will lead to better solutions, algorithms and produce high quality journal and conference publications, eventually accelerate AI and deep learning in civil engineering.

3.4. Construction and Maintenance

We can build an online dataset using the open-source data portal platform CKAN that allows easy data storage, distribution and share. CKAN is being used by public institutions [67] and government data catalogues, such as Data.gov and HealthData.gov in the US, data.gov.uk in the UK, and many others [68]. Construction and maintenance of such a dataset will need support from research grants, and donations.

4. Building and Testing a Pilot Study Database

A small dataset for concrete crack detection was built with 1499 concrete-crack images and 589 concrete-not-crack images. **Figure 2** shows the sprite image of the proposed database for concrete crack detection. For bicycle detection and counting, a database was created with 988 test images and 4822 train images. The bicycle images were taken from Google using a special data scrapping software tool and the images were labeled and annotated using LabelImg software. Moreover, a dataset was created for pavement crack detection using a 336-test image and 2284 train. The images were collected using a hand-held mobile phone and a drone. A total of 11 categories of flexible pavement crack images and 7 types of rigid pavement crack images were included in this dataset. The database images were annotated and labeled using LabelImg software with more than 50 hours of manual labors. Additionally, an infrared thermography dataset

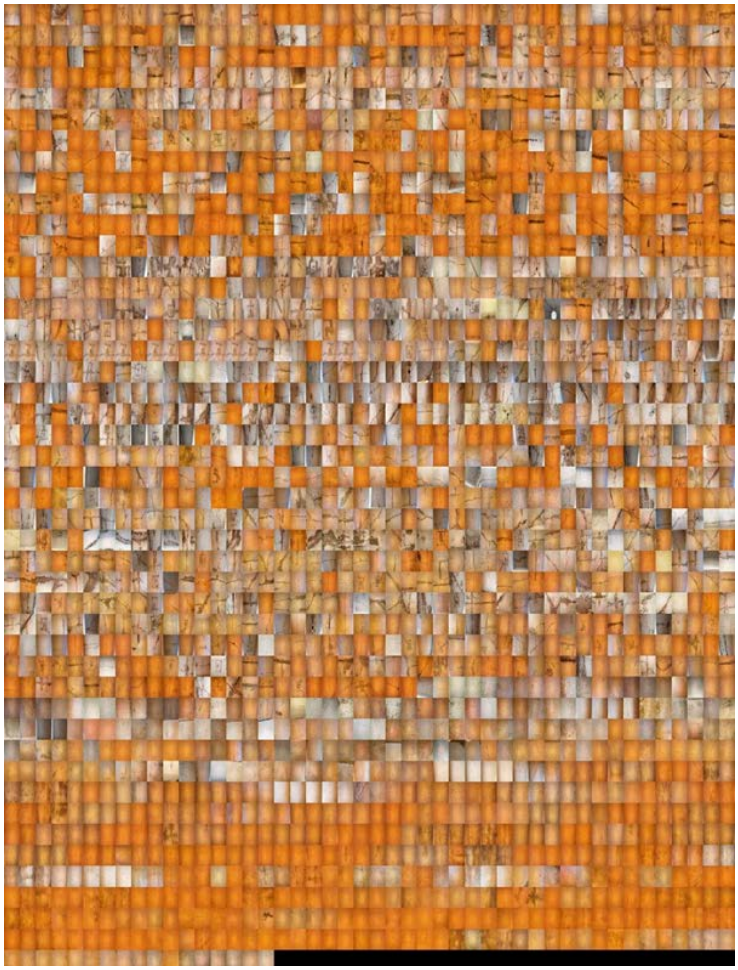


Figure 2. Sprite image of 1642 images of crack and not crack database.

was created with 24 test and 84 train infrared thermography images **Figure 3** shows the sprite images of the pavement crack detection images with infrared thermography images. The databases were uploaded in a Google share drive (http://bit.ly/2ujAhMd).

All the developed databases were tested using a deep learning convolution neural network model called Faster RCNN. The description of the test parameters and the procedures are out of the scope of this paper. The reader can find more depth in-formation of the model selection and training in this article [7]. However, the Faster RCNN model successfully detects the pavement crack using normal images (**Figure 4**) and infrared thermography crack images (**Figure 5**), as well as the pedestrian and bicyclist images (**Figure 6**). All the test images from the proposed pilot study databases show a 98% confidence level which means the database annotation and labeling are in the right direction.

5. Conclusion and Discussion

Big dataset is like fuel to engine that delivers power to the civil engineering AI research plane. A publicly available, free and labelled dataset is to address a

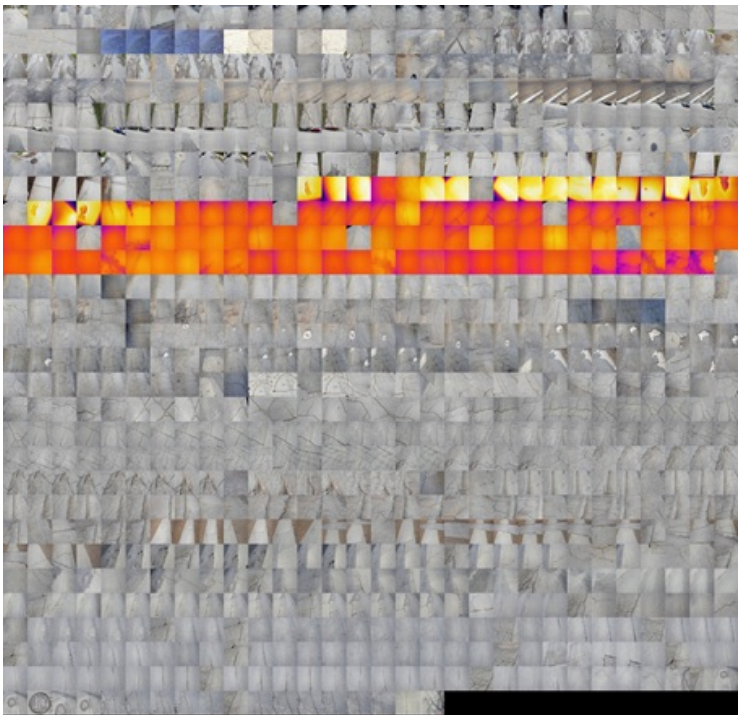


Figure 3. Sprite images of 2728 image data for pavement crack detection including infrared and normal images.

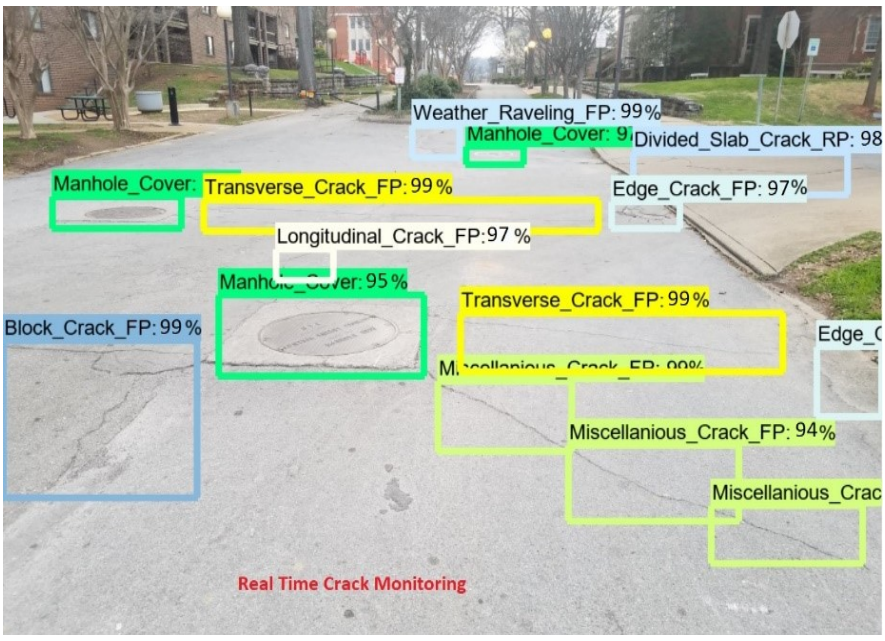


Figure 4. Pavement crack detection using Faster RCNN with 98% confidence.

fundamental issue of advancing deep learning research and application in civil engineering and beyond: high quality data. It would be a tremendous help to civil researchers to build their innovative and cutting-edge works in intelligent transportation, connected vehicle, structural health monitoring, bridge inspection, and more real-life applications on the proposed dataset. Our pilot study

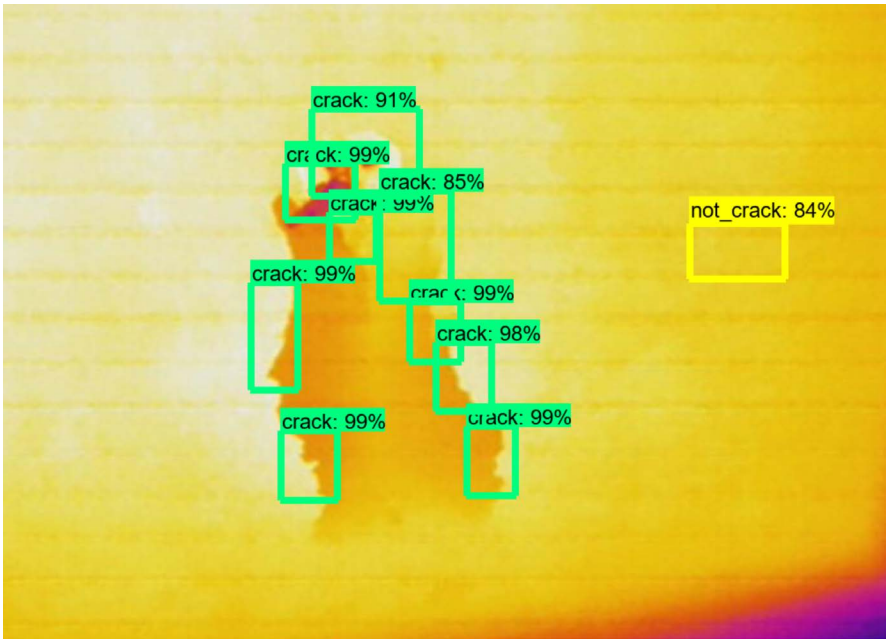


Figure 5. Concrete crack detection using faster RCNN with 98% confidence.

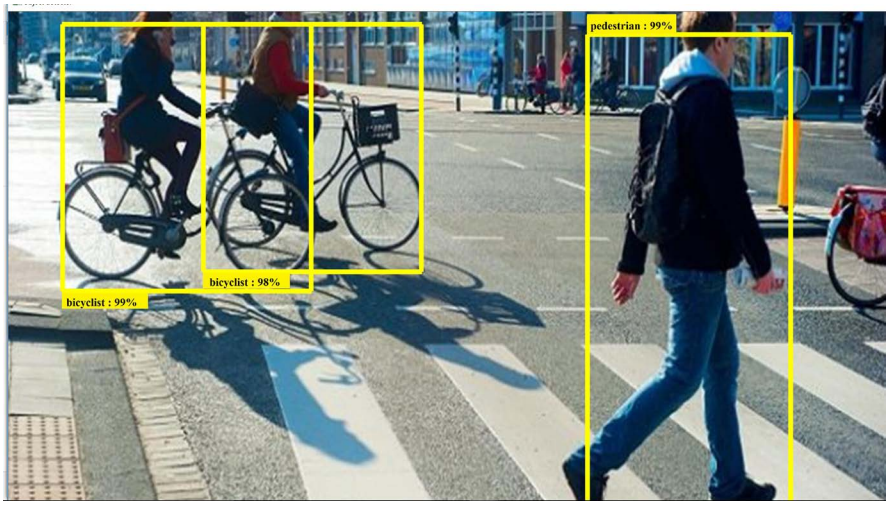


Figure 6. Pedestrian and bicyclist detection using faster RCNN with 98% confidence.

shows some of the proposed datasets for concrete crack detections, pavement crack detections as well as pedestrian and bicyclist detections with 98% confidence level.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Sato, K., Young, C. and Patterson, D. (2017) An In-Depth Look at Google's First Tensor Processing Unit (TPU). Google Cloud Big Data and Machine Learning Blog,

12.

[2] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. (2009) ImageNet: A Large-Scale Hierarchical Image Database. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, 20-25 June 2009, 248-255. https://doi.org/10.1109/CVPR.2009.5206848

[3] Chollet, F. (2017) Deep Learning with Python. Manning Publications Co., New York.

[4] Lin, Y.Z., Nie, Z.H. and Ma, H.W. (2017) Structural Damage Detection with Automatic Feature-Extraction through Deep Learning. *Computer-Aided Civil and Infrastructure Engineering*, **32**, 1025-1046. https://doi.org/10.1111/mice.12313

[5] Gulgec, N.S., Takáč, M. and Pakzad, S.N. (2017) Structural Damage Detection Using Convolutional Neural Networks. In: *Model Validation and Uncertainty Quantification*, Volume 3, Springer, New York, 331-337. https://doi.org/10.1007/978-3-319-54858-6_33

[6] Cha, Y.J., Choi, W. and Büyüköztürk, O. (2017) Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*, **32**, 361-378. https://doi.org/10.1111/mice.12263

[7] Qurishee, M.A. (2019) Low-Cost Deep Learning UAV and Raspberry Pi Solution to Real Time Pavement Condition Assessment.

[8] Cha, Y.J., Choi, W., Suh, G., Mahmoudkhani, S. and Büyüköztürk, O. (2017) Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types. *Computer-Aided Civil and Infrastructure Engineering*, **33**, 731-747. https://doi.org/10.1111/mice.12334

[9] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T. and Omata, H. (2018) Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone. arXiv Preprint arXiv:1801.09454.

[10] Makantasis, K., Protopapadakis, E., Doulamis, A., Doulamis, N. and Loupos, C. (2015) Deep Convolutional Neural Networks for Efficient Vision Based Tunnel Inspection. 2015 *IEEE International Conference on Intelligent Computer Communication and Processing* (*ICCP*), Cluj-Napoca, Romania, 3-5 September 2015, 335-342. https://doi.org/10.1109/ICCP.2015.7312681

[11] Varghese, A., Gubbi, J., Sharma, H. and Balamuralidhar, P. (2017) Power Infrastructure Monitoring and Damage Detection Using Drone Captured Images. 2017 *International Joint Conference on Neural Networks* (*IJCNN*), Anchorage, AK, 14-19 May 2017, 1681-1687. https://doi.org/10.1109/IJCNN.2017.7966053

[12] Wang, F., Kerekes, J.P., Xu, Z. and Wang, Y. (2018) Residential Roof Condition Assessment System Using Deep Learning. *Journal of Applied Remote Sensing*, **12**, Article ID: 016040. https://doi.org/10.1117/1.JRS.12.016040

[13] Russakovsky, O., *et al*. (2015) ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, **115**, 211-252. https://doi.org/10.1007/s11263-015-0816-y

[14] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. In: *Advances in Neural Information Processing Systems*, Springer, New York, 1097-1105.

[15] Lin, T.-Y., *et al*. (2014) Microsoft Coco: Common Objects in Context. In: *European Conference on Computer Vision*, Springer, New York, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48

[16] Krizhevsky, A., Nair, V. and Hinton, G. (2014) The CIFAR-10 Dataset.

https://www.cs.toronto.edu/~kriz/cifar.html

[17] Torralba, A., Fergus, R. and Freeman, W.T. (2008) 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**, 1958-1970. https://doi.org/10.1109/TPAMI.2008.128

[18] Y. LeCun, C. Cortes, and C. Burges (2010) MNIST Handwritten Digit Database. AT & T Labs. Volume 2. http://yann.lecun.com/exdb/mnist

[19] Knight, W. (2018) The White House Promises to Release Government Data to Fuel the AI Boom. https://www.technologyreview.com/s/611331/the-white-house-promises-to-release-government-data-to-fuel-the-ai-boom/

[20] Sun, C., Shrivastava, A., Singh, S. and Gupta, A. (2017) Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. 2017 *IEEE International Conference on Computer Vision* (*ICCV*), Venice, Italy, 22-29 October 2017, 843-852. https://doi.org/10.1109/ICCV.2017.97

[21] Downey, A.S. and Olson, S. (2013) Sharing Clinical Research Data: Workshop Summary. National Academies Press, Washington DC.

[22] Chollet, F. (2016) Building Powerful Image Classification Models Using Very Little Data. Volume 13.

[23] Yosinski, J., Clune, J., Bengio, Y. and Lipson, H. (2014) How Transferable Are Features in Deep Neural Networks? In: *Advances in Neural Information Processing Systems*, Springer, New York, 3320-3328.

[24] Chourabi, H., *et al*. (2012) Understanding Smart Cities: An Integrative Framework. 2012 45*th Hawaii International Conference on System Science* (*HICSS*), Maui, HI, 4-7 January 2012, 2289-2297. https://doi.org/10.1109/HICSS.2012.615

[25] Figueiredo, L., Jesus, I., Machado, J.T., Ferreira, J.R. and De Carvalho, J.M. (2001) Towards the Development of Intelligent Transportation Systems. *Intelligent Transportation Systems*, 2001. *Proceedings*, Oakland, CA, 25-29 August 2001, 1206-1211. https://doi.org/10.1109/ITSC.2001.948835

[26] Dimitrakopoulos, G. and Demestichas, P. (2010) Intelligent Transportation Systems. *IEEE Vehicular Technology Magazine*, **5**, 77-84. https://doi.org/10.1109/MVT.2009.935537

[27] Adeli, H. (2008) Smart Structures and Building Automation in the 21st Century. *International Symposium on Automation in Construction*, **25**, 5-10. https://doi.org/10.3846/isarc.20080626.5

[28] Udd, E. (1993) Fiber Optic Smart Structures. *Society of Photo-Optical Instrumentation Engineers* (*SPIE*) *Conference Series*, Volume 10266. https://doi.org/10.1117/12.145196

[29] Stajano, F., Hoult, N., Wassell, I., Bennett, P., Middleton, C. and Soga, K. (2010) Smart Bridges, Smart Tunnels: Transforming Wireless Sensor Networks from Research Prototypes into Robust Engineering Infrastructure. *Ad Hoc Networks*, **8**, 872-888. https://doi.org/10.1016/j.adhoc.2010.04.002

[30] Lajnef, N., Chatti, K., Chakrabartty, S., Rhimi, M. and Sarkar, P. (2013) Smart Pavement Monitoring System. United States Federal Highway Administration.

[31] Barbaresso, J., Cordahi, G., Garcia, D., Hill, C., Jendzejec, A. and Wright, K. (2014) USDOT's Intelligent Transportation Systems (ITS) ITS Strategic Plan 2015-2019. https://doi.org/10.1109/JPROC.2011.2132790

[32] Kenney, J.B. (2011) Dedicated Short-Range Communications (DSRC) Standards in

the United States. *Proceedings of the IEEE*, **99**, 1162-1182.

[33] Hartman, K.K. Connected Vehicle Pilots. https://www.its.dot.gov/pilots/index.htm

[34] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779-788. https://doi.org/10.1109/CVPR.2016.91

[35] Ren, S., He, K., Girshick, R. and Sun, J. (2015) Faster r-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: *Advances in Neural Information Processing Systems*, Springer, New York, 91-99.

[36] Farrar, C.R. and Worden, K. (2007) An Introduction to Structural Health Monitoring. *Philosophical Transactions of the Royal Society of London A*: *Mathematical, Physical and Engineering Sciences*, **365**, 303-315. https://doi.org/10.1098/rsta.2006.1928

[37] Kim, S., *et al*. (2007) Health Monitoring of Civil Infrastructures Using Wireless Sensor Networks. In: *Proceedings of the* 6*th International Conference on Information Processing in Sensor Networks*, ACM, New York, 254-263. https://doi.org/10.1145/1236360.1236395

[38] Wu, W., Qurishee, M.A., Owino, J., Fomunung, I., Onyango, M. and Atolagbe, B. (2018) Coupling Deep Learning and UAV for Infrastructure Condition Assessment Automation. 2018 *IEEE International Smart Cities Conference* (*ISC*2), Kansas City, MO, 16-19 September 2018, 1-7. https://doi.org/10.1109/ISC2.2018.8656971

[39] Polson, N.G. and Sokolov, V.O. (2017) Deep Learning for Short-Term Traffic Flow Prediction. *Transportation Research Part C: Emerging Technologies*, **79**, 1-17. https://doi.org/10.1016/j.trc.2017.02.024

[40] Lv, Y., Duan, Y., Kang, W., Li, Z. and Wang, F.-Y. (2015) Traffic Flow Prediction with Big Data: A Deep Learning Approach. *IEEE Transactions on Intelligent Transportation Systems*, **16**, 865-873.

[41] Mocanu, E., Nguyen, P.H., Gibescu, M. and Kling, W.L. (2016) Deep Learning for Estimating Building Energy Consumption. *Sustainable Energy*, *Grids and Networks*, **6**, 91-99. https://doi.org/10.1016/j.segan.2016.02.005

[42] Li, C., Ding, Z., Zhao, D., Yi, J. and Zhang, G. (2017) Building Energy Consumption Prediction: An Extreme Deep Learning Approach. *Energies*, **10**, 1525. https://doi.org/10.3390/en10101525

[43] NHTS Administration (2017) Federal Motor Vehicle Safety Standards; V2V Communications. *Federal Register*, **82**, 3854-4019.

[44] Miller, G.A. (1995) WordNet: A Lexical Database for English. *Communications of the ACM*, **38**, 39-41. https://doi.org/10.1145/219717.219748

[45] Ryan, T.W., Hartle, R.A., Mann, J.E. and Danovich, L.J. (2006) Bridge Inspector's Reference Manual. Report No. FHWA NHI, 03-001.

[46] Russell, B.C., Torralba, A., Murphy, K.P. and Freeman, W.T. (2008) LabelMe: A Database and Web-Based Tool for Image Annotation. *International Journal of Computer Vision*, **77**, 157-173. https://doi.org/10.1007/s11263-007-0090-8

[47] Von Ahn, L. and Dabbish, L. (2004) Labeling Images with a Computer Game. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, 319-326. https://doi.org/10.1145/985692.985733

[48] Benfield, J.A. and Szlemko, W.J. (2006) Internet-Based Data Collection: Promises and Realities. *Journal of Research Practice*, **2**, 1.

[49] Kramer, V.H. and Weinberg, D.B. (1974) The Freedom of Information Act. *The*

Georgetown Law Journal, **63**, 49.

[50] Caltrans. Performance Measurement System (PeMS). http://pems.dot.ca.gov

[51] Dot, T. Smartway. https://smartway.tn.gov/traffic/

[52] GDOT. The Georgia Department of Transportation's Traffic Analysis and Data Application. https://gdottrafficdata.drakewell.com/publicmultinodemap.asp

[53] UDO Transportation. ITS Public Data Hub. https://www.its.dot.gov/data/

[54] Qurishee, M., Iqbal, I., Islam, M. and Islam, M. (2016) Use of Slag as Coarse Aggregate and Its Effect on Mechanical Properties of Concrete. *Proceedings of International Conference on Advances in Civil Engineering*, **3**, 475-479.

[55] Al Qurishee, M. (2017) Application of Geosynthetics in Pavement Design.

[56] Al Qurishee, M. and Fomunung, I. (2000) Smart Materials in Smart Structural Systems.

[57] Hasnat, A., Qurishee, M., Iqbal, I., Zaman, M. and Wahid, M. (2018) Effectiveness of Using Slag as Coarse Aggregate and Study of Its Impact on Mechanical Properties of Concrete.

[58] Atolagbe, B. (2019) Automatic Mesh Representation of Urban Environments.

[59] Islam, M.A. (2018) Intergrading Connected Vehicle Data into the Transportation Performance Measurement Process. The University of Alabama, Birmingham.

[60] Islam, M.A. (2019) A Literature Review on Freeway Traffic Incidents and Their Impact on Traffic Operations. *Journal of Transportation Technologies*, **9**, 504-516. https://doi.org/10.4236/jtts.2019.94032

[61] Islam, M.A., Sisiopiku, V.P., Ramadan, O.E. and Hadi, M. (2019) A Framework for Performance-Based Traffic Operations Using Connected Vehicle Data. *Simulation* (*NGSIM*), **6**, No. 8.

[62] Al Qurishee, M., Wu, W., Atolagbe, B., El Said, S. and Ghasemi, A. (2019) Non-Destructive Test Application in Civil Infrastructure.

[63] Al Qurishee, M., Wu, W., Atolagbe, B., El Said, S., Ghasemi, A. and Tareq, S.M. (2008) Wireless Sensor Network and Its Application in Civil Infrastructure.

[64] Cdot (2020) California Department of Transportation. http://pems.dot.ca.gov/

[65] Tdot (2020) SmartWay Traffic. https://smartway.tn.gov/Traffic?position=-85.97076699999997,35.88416963357008,7&features=incidents,traffic

[66] Usdot (2020) ITS DataHub. https://www.its.dot.gov/data/

[67] Winn, J. (2013) Open Data and the Academy: An Evaluation of CKAN for Research Data Management.

[68] OK International. https://ckan.org/

Scientific Research Publishing

Engineering, 2020, 12, 166-176
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

Detection of "Swollen Shoot" Disease in Ivorian Cocoa Trees via Convolutional Neural Networks

Mamadou Coulibaly, Konan Hyacinthe Kouassi, Silue Kolo, Olivier Asseu

Ecole Supérieure Africaine des TIC, LASTIC, Abidjan, Côte d'Ivoire
Email: oasseu@yahoo.fr

How to cite this paper: Coulibaly, M., Kouassi, K.H., Kolo, S. and Asseu, O. (2020) Detection of "Swollen Shoot" Disease in Ivorian Cocoa Trees via Convolutional Neural Networks. *Engineering*, **12**, 166-176.
https://doi.org/10.4236/eng.2020.123014

Received: February 13, 2020
Accepted: March 9, 2020
Published: March 12, 2020

Abstract

Recent advances in diagnostics have made image analysis one of the main areas of research and development. Selecting and calculating these characteristics of a disease is a difficult task. Among deep learning techniques, deep convolutional neural networks are actively used for image analysis. This includes areas of application such as segmentation, anomaly detection, disease classification, computer-aided diagnosis. The objective which we aim in this article is to extract information in an effective way for a better diagnosis of the plants attending the disease of "swollen shoot".

Keywords

Drone, Convolutional Neural Networks, Image Recognition, Feature Detection

1. Introduction

The Ivorian cocoa farm is in danger. A disease spreads quickly in the Ivorian cocoa farm. This is the "Swollen shoot", a viral disease first described in Ghana in 1936 [1]. This disease, largely confined to West Africa, is considered a major viral disease. It has caused invaluable losses to producers in the Ivory Coast and Ghana where several million cocoa plants have been destroyed [1]. The virus is mainly transmitted by mealybugs. Three main symptoms are indeed associated with the development of the disease: swelling of twigs and roots, yellowing of leaves and deformation of organs.

The fight against the "Swollen shoot" currently consists of circumscribing it, since there is to date no cure for this virus [1]. This process is carried out on 4 components: prospecting and delimiting infected areas; uprooting and destruction of infected feet; replanting with a minimum distance of 10 m between old

and new plantations. But this method of prospecting remains rudimentary. Since you have to go there and the risks of physical exposure to this scourge are not negligible.

To contribute effectively to the fight against this epidemic we decided to set up a system of recognition of the symptoms of "Swollen shoot" by drone, based on the networks of convolutional neurons.

Côte d'Ivoire is a country whose agriculture accounts for more than a third of the Gross Domestic Product (GDP). Agriculture is an area in which almost all artisans are strongly involved, more specifically for cocoa growing. The epidemic of "Swollen shoot" requires that these artisans know the state of health of their plant. Moving an agronomist every time a plant has anomalies would be a waste of time and cost the planter enormous costs. It would be better for them to have an immediate solution to diagnose their plants at any time and at very low cost. It is in this context that we decided to set up a system based on convolutional neural networks for the detection of "Swollen shoot".

2. Convolutional Neural Networks

Deep learning is a tool used for machine learning, in which several linear and non-linear processing units are arranged in a deep architecture to model the high level abstraction present in the data [2]. There are many deep learning techniques currently used in various applications. These include auto-encoders, stacked auto-encoders, restricted Boltzmann machines (RBM), deep belief networks (DBN) and deep convolutional neural networks (CNN). In recent years, CNN-based methods have gained popularity in vision systems as well as in the field of medical image analysis [3] [4] [5].

CNNs combine three architectural ideas to ensure, to some extent, the invariance in terms of scale, offset and distortion. The first CNN model (LeNet-5) proposed for the recognition of handwritten characters is presented in [6]. The local pattern connections between the neurons of the adjacent CNN layers, *i.e.* the inputs of the hidden units of a layer m, are considered to be a subset of units of the layer $m-1$, units with adjacent receiving fields to exploit local spatial correlation. In addition, in CNN, each h_l filter is replicated across the entire visual field. These filters share weighting and bias vectors to create a feature map. The gradient of the shared weights is equal to the sum of the gradients of the shared parameters. When the convolution operation is performed on sub-regions of the entire image, a characteristic map is obtained.

The process involves the convolution of the input image or the characteristic map with a linear filter, with the addition of a bias followed by the application of a non-linear filter. A bias value is added so that it is independent of the output of the previous layer. The bias values allow us to shift the activation function of a node to the left or to the right. For example, for a sigmoid function, the weights control the inclination of the output, while the bias is used to compensate for the curve and allows a better fit of the model. Bias values are learned during the

training model and allow an independent variable to control activation. On a given layer, the k^{th} filter is symbolically designated by h^k, and the weights W^k and bias b_k determine their filters. The mathematical expression for obtaining characteristic maps is given as follows:

$$h_{ij}^k = \tanh\left(\left(W^k * x\right)_{ij} + b_k\right) \tag{1}$$

where, tanh represents the tan hyperbolic function, and $*$ is used for the convolution operation. Figure 1 illustrates two hidden layers in a CNN, where layer $m-1$ and m have four and two features maps respectively *i.e.*, h^0 and h^1 named as w^1 and w^2. These are calculated from pixels (neurons) of layer $m-1$ by using a 2×2 window in the layer below as shown in Figure 1 by the colored squares. The weights of these filter maps are 3D tensors, where one dimension gives indices for input feature maps, while the other two dimensions provide pixel coordinates. Combining it all together, W_{ij}^{kl} represents the weight connected to each pixel of kth feature map at a hidden layer m with ith feature map of a hidden layer $m-1$ and having coordinates i, j.

Each neuron or node in a deep network is governed by an activation function that controls the output. There are various activation functions used in the deep learning literature, such as linear, sigmoid, tanh, rectified linear unit (ReLU).

The neural network that we offer contains: a convolution layer, a pooling layer, a correction layer and a fully-connected layer.

- The convolution layer is the key component of CNNs, and always constitutes at least their first layer. The convolution layer therefore receives as input several images, and calculates the convolution of each of them with each filter. The filters correspond exactly to the characteristics that one wishes to find in the images.

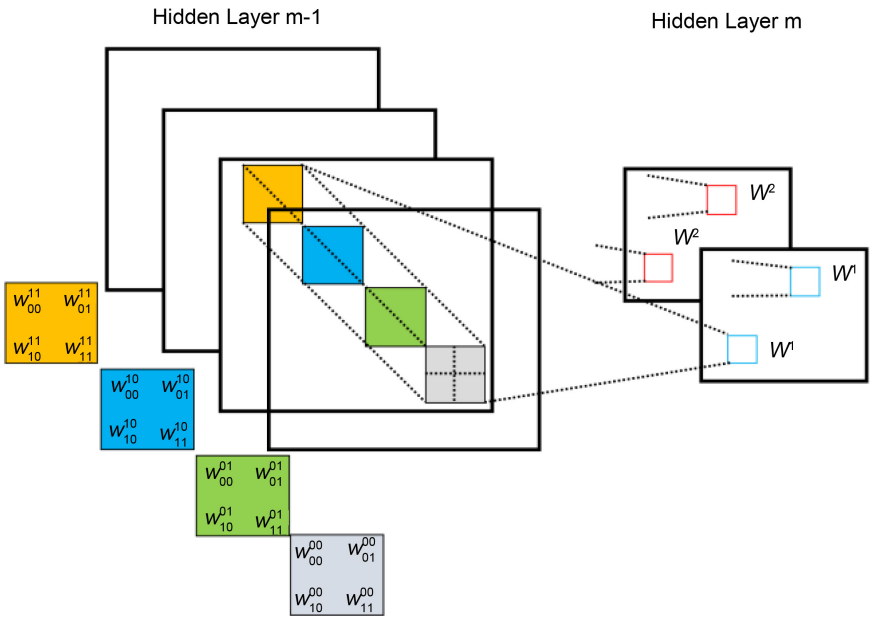


Figure 1. Hidden layers in a convolutional neural network.

- The pooling layer is this type of layer is often placed between two convolutional layers: it receives as input several characteristics, and applies the pooling operation to each of them. The pooling operation consists in reducing the size of the images, while preserving their important characteristics. For this, we cut the image into regular cells, then we keep within each cell the maximum value.
- The ReLU (Rectified Linear Units) correction layer designates the real non-linear function defined by: ReLU(x) = max(0, x). The ReLU correction layer therefore replaces all negative values received as inputs with zeros. It plays the role of activation function.
- The last fully-connected layer makes it possible to classify the image at the input of the network: it returns a vector of size N, where N is the number of classes in our image classification problem. Each element of the vector indicates the probability for the input image to belong to a class. For example, if the problem consists in distinguishing cocoa in good condition and cocoa infected with Swollen, the final vector will be of size 2: the first element (respectively, the second) gives the probability of belonging to the class "healthy cocoa" (respectively "cocoa infects"). Thus, the vector [0.90 0.1] means that the image has a 90% chance of representing cocoa in good condition.

Below is the model of our neural network (Figure 2).

3. Presentation of the System and Data Analysis

The system we offer is made up of four main phases (Figure 3):

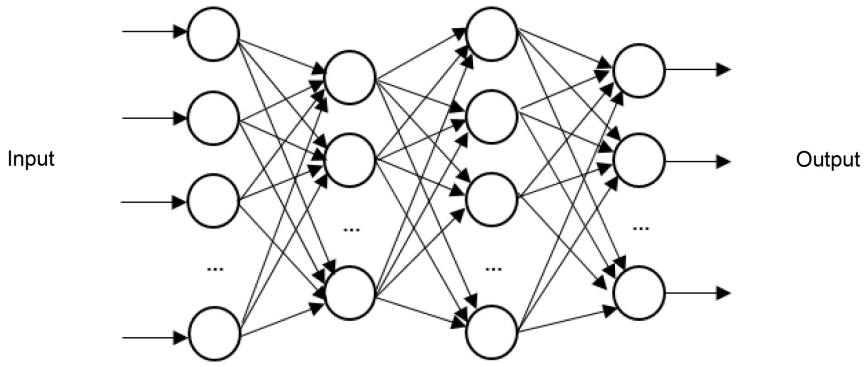


Figure 2. Convolutional neural network.

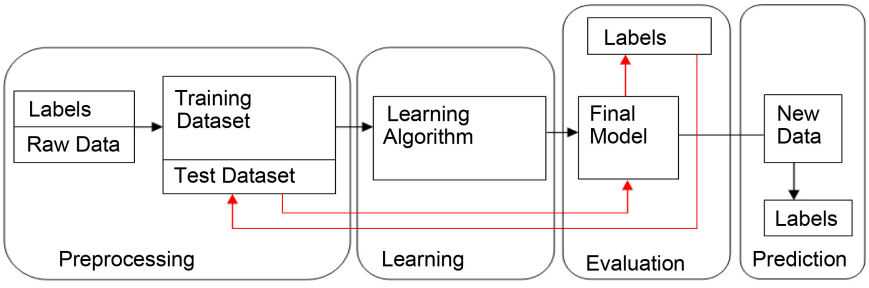


Figure 3. Modeling our system.

- The pretreatment phase;
- The learning phase;
- The evaluation phase;
- The prediction phase.

The first thing to do is to explore all possible avenues to recover the input data from our system. Indeed, data constitutes experience, the examples that we will provide to your algorithm so that it can learn and become more efficient. Once the data has been found, you must proceed to the cleaning step. Care should be taken to ensure that the data is consistent, with no outliers or missing values. We have few datasets for the detection of symptoms of "Swollen shoot". To have more data, we applied geometric transformations, which modify the position of the pixels in the image. A rotation, a translation or a change of scale (zoom) are typical examples.

4. Data Preprocessing

Two images of the same class, which represent the same physical object, are therefore linked by a transformation. However, the precise characterization of this transformation is unknown to us. How to determine the transformation which makes it possible to pass from an image to another? It is in fact a classic problem in computer vision, called image matching. Instead of trying to determine the precise mathematical equation of the transformation as in the previous part, the strategy consists in finding the elements common to the two images called patterns. These patterns are represented by small images, called Template. The purpose of Template matching is to find patterns in an image.

The template matching carried out with filters uses the cross-correlation operator, noted $\otimes$. This operator transforms the image of matrix representation X into a new image $Y = H \otimes X$ in the following way:

$$Y_{i,j} = \sum_{u=-k}^{k} \sum_{v=-k}^{k} H_{u,v} X_{i+u,j+v} \tag{2}$$

In this context, H is a small image representing the Template to find. Concretely, this operation amounts to dragging H onto the image X, to multiply the overlapping pixels and to summing these products.

Thus, Template matching consists in calculating the cross correlation between an image X and a filter whose kernel H represents a Template that one wishes to find in X. In reality, this technique is not so practical. On the one hand, to be able to define the Template, we have to look for the image area to recognize by hand before doing template matching. On the other hand, the cross-correlation operator is very sensitive to these variations. So we have to find a way to get a more generic template. In other words, it is necessary to formalize the properties of the characteristic elements of a class of images. This is where the notion of characteristic of an image comes in.

In computer vision, the term features designate characteristic areas of the digital image. These areas can correspond to outlines, points or regions of interest. Each detected feature is associated with a vector, called a descriptor (feature de-

scriptor or feature vector), which, as its name suggests, describes the area concerned. The resolution of the image matching problem is then done in two stages (Figure 4):

1) Detect and describe the features in each image.

2) Find the pairs of features that match in the two images (features matching).

The image matching algorithm studies the characteristics of the images, so the quality of the results depends (among other things) on the relevance of the detected characteristics. In this sense, the first step is fundamental and should in no case be overlooked.

A wrong choice of characteristics can lead to several difficulties in the matching stage:

- Problem 1: two images do not have the same characteristics when they represent the same object in different ways.
- Problem 2: these two images have the same characteristics, but finding the matching pairs is very difficult.

These two problems make the correspondence impossible and must therefore be anticipated from the first step, when detecting and describing characteristics. This brings us to the following question:

What characteristics should be selected?

We have introduced the features as interesting areas of the image. More specifically, an area is a good choice of characteristics if it is:

- Repeatable: a characteristic must be found in the images representing the same object despite the geometric and photometric differences. A characteristic must therefore have invariance properties to these transformations.
- Distinctive: a characteristic must be sufficiently unique and unambiguous within an image to facilitate correspondence. It is the information contained in its descriptor which must highlight its particularity.
- Local: a characteristic must correspond to a sufficiently small zone, and it is described according to its vicinity only. This avoids the difficulties of correspondence due to occlusion phenomena.

The gradient of an image is a very useful tool for characteristic detection. It is a vector ∇I composed of the partial derivatives of the intensity function, and calculated in each pixel:

Figure 4. Legend example of image matching. The characteristics are marked by the squares.

$$\nabla I(x,y) = \left(\frac{\partial I(x,y)}{\partial x}, \frac{\partial I(x,y)}{\partial y} \right) \tag{3}$$

The partial derivative with respect to x (or y) makes it possible to study the variations in image intensities in the direction of the abscissa (or ordinate) axis.

5. Network Learning

After having carried out the model, it is necessary to submit it to the data so that it adjusts the values of its parameters (if it has any). The parameters are adjusted using a learning algorithm. If, in addition to presenting the learning data at the input of the algorithm, the desired outputs are presented, then learning is said to be supervised. In the context of supervised learning, if the desired output takes its value in a set of finite cardinals then the task performed is a classification. Otherwise, we talk about regression. Furthermore, if the training data is submitted to the algorithm only without the desired outputs, then the learning is said to be unsupervised. In this case, the task of the algorithm is either to find interesting relationships between the data, or to partition them (clustering) according to the predefined similarity criteria. In some cases, unsupervised learning aims to reduce the dimension of observations [7].

Deep learning is a new form of machine learning that uses neural networks. Flow graphs 4 are used to represent neural networks. The adjective "deep", which comes to qualify learning or neural networks, originates from a property of flow graphs: depth [8]. The idea of deep neural networks has evolved very quickly over time. Twenty years ago, a network of more than two hidden layers was considered deep. This idea of the depth of networks is over. Today, the number of deep neural network layers is in the hundreds. Another characteristic of deep neural networks is the type of layers that make it up. New layers such as convolutional layers are specific to deep neural networks [9] [10]. In the past decade, deep learning has been the subject of particular interest in the field of artificial intelligence [11]. Several large companies have engaged in the production of deep learning technologies, including Google, Facebook, Microsoft and Yahoo.

6. Network Evaluation

Validation is a phase which evaluates the learning model. There are two levels of validation: one at the model selection level, and the other at the model parameter selection level. Indeed, several models are produced and trained. The validation phase therefore consists in selecting the model that best reflects reality first and then adjusting the parameters again. This is done using tests. It is usual to provide a database devoted to tests so as not to bias the results. To this end, several specialists suggest separating the data collected into two or three groups (bases): a training base, a validation base and possibly a test base [12].

This is a study conducted on 10,000 cocoa plants, we obtain the following confusion matrix:

	Healthy cocoa	Infected cocoa	Total
Symptom detected	800	4902	5702
No symptoms	4200	98	4298
Total	5000	5000	10,000

Call the class corresponding to a healthy cocoa tree "positive" and the other class "negative". If we detect the symptoms of the Swollen shoot when there is one, we make a "positive" prediction which is correct, it is a true positive. If, on the other hand, this prediction is incorrect, it is a false positive. And so on. Also sometimes called "type I error" false positives, and "type II error" false negatives.

We thus define "recall", or "sensitivity", is the rate of true positives, that is to say the proportion of positives that we have correctly identified. This is the ability of our model to detect all cocoa plants in good condition:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

We will also be interested in "precision", that is to say the proportion of correct predictions among the points that we predicted positive. It's our model's ability to only trigger an alarm for a real fire.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

To finish this long list, we are also often interested in "specificity", which is the rate of true negatives, in other words the ability to detect all situations where there are symptoms of Swollen shoot detected. It is a complementary measure of sensitivity.

$$\text{specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \tag{6}$$

- Recall = 84%
- Specificity = 98%
- Precision = 74%

7. Conclusion

To effectively contribute to the fight against the Swollen shoot epidemic, we have set up a symptom recognition system, based on convolutional neural networks. Our results show that a large network of deep convolutional neurons is capable of obtaining record results on a data set using supervised learning. Indeed, the probability of the system to predict that the cocoa plant is healthy when this is true in reality is 4200/5000 or 84%, which makes this test a good detection tool. It should be noted that the performance of our network degrades if a single convolutional layer is removed. So depth is really important to achieve our results. Our results can be improved if we expand and train our network longer. Ultimately, we would like to use very large and deep convolutional networks on video footage where the time structure provides very useful information that is

missing or much less evident in static images.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] CNRA (2006) Bulletin d'information et de liaison du Centre National de Recherche Agronomique N° 18 Août 2006.

[2] Deng, L., Yu, D., *et al.* (2014) Deep Learning: Methods and Applications, *Foundations and Trends R in Signal Processing*, **7**, 197-387. https://doi.org/10.1561/2000000039

[3] Premaladha, J. and Ravichandran, K. (2016) Novel Approaches for Diagnosing Melanoma Skin Lesions through Supervised and Deep Learning Algorithms. *Journal of Medical Systems*, **40**, 96. https://doi.org/10.1007/s10916-016-0460-2

[4] Kharazmi, P., Zheng, J., Lui, H., Wang, Z.J. and Lee, T.K. (2018) A Computer-Aided Decision Support System for Detection and Localization of Cutaneous Vasculature in Dermoscopy Images via Deep Feature Learning. *Journal of Medical Systems*, **42**, 33. https://doi.org/10.1007/s10916-017-0885-2

[5] Wang, S.-H., Phillips, P., Sui, Y., Liu, B., Yang, M. and Cheng, H. (2018) Classification of Alzheimers Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. *Journal of Medical Systems*, **42**, 85. https://doi.org/10.1007/s10916-017-0885-2

[6] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition, *Proceedings of the IEEE*, **86**, 2278-2324. https://doi.org/10.1109/5.726791

[7] Delalleau, O. (2012) Apprentissage machine efficace: Théorie et pratique. Thèse de doctorat, Université de Montreal, Montreal.

[8] Labelle, J. (1981) Theorie des graphes. Modulo, Montréal.

[9] Zeiler, M. and Fergus, R. (2013) Visualizing and Understanding Convolutional Networks. arXiv:1311.2901

[10] LeCun, Y. (1988) A Theoretical Framework for Back-Propagation. In: Touretzky, D., Hinton, G. and Sejnowski, T., Eds., *Proceedings of the* 1988 *Connectionist Models Summer School*, CMU, Pittsburg, PA.

[11] Copeland, M. (2006) What's the Difference between Artificial Intelligence, Machine Learning, and Deep Learning? The Canadian Press.

[12] Lerman, L. (2011) Les systemes de detection d'intrusion bases sur du machine learning.

Annexes

The training of a neural network consists in determining and empirically calculating the value of each of its parameters. The principle is as follows: the network processes an image and at the output it makes a prediction, that is to say that it tells which class it thinks this image belongs to. Knowing that we already know the class of each training image, we can check if this result is correct. Depending on the veracity of this result, we update all the network parameters, according to an algorithm called backpropagation of the error gradient. The aim of the gradient algorithm is to converge iteratively towards an optimal configuration of the synaptic weights (parameters). This state can be a local minimum of the function to be optimized and ideally a global minimum of this function, called the cost function. The weights in the neural network are first initialized with random values. We then consider a set of data that will be used for learning. Each sample has its target values which are those which the neural network must ultimately predict when presented with the same sample. The "Stochastic gradient descent" algorithm provides our trained neural network. It uses the "Backpropagation" algorithm to calculate the stochastic gradient. The two algorithms are presented below.

Algorithme 1: Backpropagation

1: Input: m pairs of i.i.d. supervised examples $\left(\tilde{x}_i, x_i\right)_{1\le i\le m}$

2: Parameter: connection weights and biases of the neural network $\theta := \left(W_l, b_l\right)_{1\le l\le n+1}$

3: Output: derivative of the objective

$$J_m = \frac{1}{2m}\sum_{i=1}^{m}\left\| f\left(\tilde{x}_i,\theta\right) - x_i\right\|_2^2$$

with respect to θ.

Begin

4: Feedforward to get a m-column matrix D_{n+1} whose ith column represents $f\left(\tilde{x}_i,\theta\right) - x_i$ as well as the m-column activation matrices $\left(A_l\right)_{1\le l\le n+1}$ whose ith column is formed by the values computed at layer l from the input $\tilde{x}_i$ according to (**2**)

5: **for** $l \leftarrow n+1$ **to** 1 **do**

6: Compute the derivatives

$$\frac{\partial J_m}{\partial b_l} = \frac{1}{m} D_l 1_{m\times 1}$$

$$\frac{\partial J_m}{\partial W_l} = \frac{1}{m} D_l A_l^{\mathrm{T}}$$

$$\frac{\partial J_m}{\partial A_l} = W_l^{\mathrm{T}} D_l$$

Where $1_{m\times 1}$ is the column vector of m ones.

7: Set 1_l to be a matrix filled with ones and of the same dimension as A_l

$$D_{l-1} = \frac{\partial J_m}{\partial A_l} \square \left(1_l - A_l \square A_l\right)$$

8: **endfor**

9: Order $\left(\frac{\partial J_m}{\partial W_l}, \frac{\partial J_m}{\partial b_l}\right)_{1\le l\le n+1}$ the same way as $\left(W_l, b_l\right)_{1\le l\le n+1}$ in θ to form $\frac{\partial J_m}{\partial \theta}$

End

Algorithme 2: Stochastic gradient descent

1: Input: initial parameter θ_0 a training dataset and a distinct validation set $(\tilde{x}_i, x_i)_{1\le i\le N_1}$

2: Parameter: learning rate α, iteration number T, observation interval S, batch size N_2

3: Output: a trained neural network θ_*.

Begin

4: Evaluate the current error on the validation set

$$\grave{o} := \frac{1}{N_1}\sum_{i=1}^{N_1}\left\| f(\tilde{x}_i, \theta_0) - x_i \right\|_2^2$$

5: **for** $t = 0$ to $T-1$ **do**

6: Draw N_2 pairs of supervised examples from the training dataset.

7: Calculate the stochastic gradient with Algorithm 1:

$$\frac{\partial}{\partial\theta} J_{N_2}(\theta_t)$$

8: Update the neural network

$$\theta_{t+1} = \theta_t - \alpha\frac{\partial}{\partial\theta} J_{N_2}(\theta_t)$$

Note that one may also want to set the learning rate in a layer-wise fashion, in which case α is a positive diagonal matrix.

9: **if** mod $(t+1, S) = 0$ **then**

10: Evaluate the current error on the validation set

$$\grave{o} := \frac{1}{N_1}\sum_{i=1}^{N_1}\left\| f(\tilde{x}_i, \theta_{t+1}) - x_i \right\|_2^2$$

11: **if** $\grave{o} > e$ **then**

12: $\grave{o} \leftarrow e$ and $\theta_* \leftarrow \theta_{t+1}$

13: end if

14: end if

15: end for

Scientific Research Publishing

Engineering, 2020, 12, 177-193
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

Finite Element Modelling of Car Seat with Hyperelastic and Viscoelastic Foam Material Properties to Assess Vertical Vibration in Terms of Acceleration

Purnendu Mondal, Subramaniam Arunachalam

School of Architecture, Computing and Engineering, University of East London (Docklands Campus), London, UK
Email: u1619864@uel.ac.uk, s.arunachalam@uel.ac.uk

How to cite this paper: Mondal, P. and Arunachalam, S. (2020) Finite Element Modelling of Car Seat with Hyperelastic and Viscoelastic Foam Material Properties to Assess Vertical Vibration in Terms of Acceleration. *Engineering*, **12**, 177-193.
https://doi.org/10.4236/eng.2020.123015

Received: February 16, 2020
Accepted: March 20, 2020
Published: March 23, 2020

Abstract

Primary objective of automobile seats is to offer adequate level of safety and comfort to the seated human occupant, primarily against vibration. Ideally, any sort of automotive seat is constructed by mechanical framework, cushion, backrest and headrest. The frame structures are made of metallic alloys, while the cushion, backrest and headrest are made of polyurethane foam material. During the design phase of automotive seat, the greatest challenge is to assign realistic material properties to foam material; as it is non-linear in nature and exhibit hysteresis at low level stress. In this research paper, a car seat has been modelled in finite element environment by implementing both hyperelastic and viscoelastic material properties to polyurethane foam. The car seat has been excited with the loads due to car acceleration and human object and the effects of vibration in terms of vertical acceleration at different locations have been measured. The aims of this simulation study are to establish a car seat with the foam material properties as accurately as possible and provide a finite element set up of car seat to monitor the vertical acceleration responses in a reasonable way. The RMS acceleration values for headrest, backrest and cushion have been found to be 0.91 mm/sec^2, 0.54 mm/sec^2 and 0.47 mm/sec^2, respectively, which showed that the car seat foam can effectively be modelled through combined hyperelastic and viscoelastic material formulations. The simulation outputs have been validated through real life testing data, which clearly indicates that this computerized simulation technique is capable of anticipating the acceleration responses at different car seat segments in a justified way.

Keywords

Car Seat, Hyperelastic Material, Viscoelastic Material, Finite Element,

Vertical Acceleration, Vertical Vibration

1. Introduction

The car seat foam can be made of nylon, alcantara, vinyl, faux leather or polyester material, though the most common industrially used material for car seat foam is polyurethane. Over the last many years, numerous research studies had been carried out on car seat based on the shape, orientation, material to optimize the human safety and comfort standard, though seat material properties were always accounted as the mandatory factor regardless of the field of research.

Mechanical behaviour of polyurethane material is highly complex in nature. For a small quantity of strain effect, it continues to behave like elastic material, though further increment in strain amount causes gradual increment of stress generated inside it. At the end of compression stage, it exhibits sharp rise in stress level. Non-linearity with hysteresis and strain rate dependent energy dissipative nature of polyurethane foam material were shown in the finite element analysis of car seat [1] and load deflection measurement of foam material [2]. Simulation using LS DYNA for obtaining the force-deflection curve for the seat cushion had been performed [3] to show the static and dynamic characteristics of the cushion foam material. The seat base plate, the seat cushion and a circular disc had been taken into account to carry out the entire analysis and later the results were compared to experimental data. The seat cushion material properties had been simplified by assigning only density and Young's modulus. The influence of the seat foam material properties had been shown during the assessment of sitting comfort of the automobile seats [4] for contact interference and acceleration transmission between human body and seat. Various physical parameters including seat material properties had been explored [5] during the pressure distribution study between seat and sitting object. Driver's seat comfort was virtually simulated in finite element [6] to observe the pressure distribution inside seat with respect to inclination of the seat. Simplified seat model had been constructed, made of backrest and cushion with assigned density, Poisson's ratio and Young's modulus. Human health and safety had been investigated during the course of collision [7] and found that seat material would be responsible for the amount of movement of human occupant in the side and rear directions. Experimental and finite element studies had been conducted on polyurethane foam [8] to show the compression characteristics under the effect of random vibration. Researchers established a series of feasible models for foam material including hyperelastic or hyper-foam constitutive model. Co-relations between the stress, strain and volume of the foam material had been formulated while studying foam properties in depth [9]. Many research works in the past attempted to establish the properties of the polyurethane foam by relating the stress-strain behaviour as function of strain rate. Finite element investigation on car seat [1]

stated that the stress-strain curve for foam material was directly related to strain rate, while study on foam material under the effect of dynamic impact load [10] graphically represented the stress-strain curve with respect to various uniaxial compressions with variable strain rates. Seat for heavy vehicle had been numerically simulated using lumped network system modelling [11] to monitor the riding comfort. The seat foam material had been modelled using both the hyper-elastic and viscoelastic material parameters to take into account the strain rate dependency and strain energy dissipation. Vibration transmission from seat leg to passenger body had been observed [12] and vibration damping in passenger seat had been assessed. Open-porous aluminum foam material had been considered and modal damping state followed by frequency response function and acceleration graph had been received. Seat cushion simulation under the effects of uniaxial compression and indentation forces [13] found that modelling the foam material with hyper-foam formulation was capable of anticipating the stress-strain curve in an accurate way. Finite element modelling of the human soft tissue and seat foam [14] used the hypothetical concept of strain energy absorption capacity along with hyperelastic material properties as described in the Ogden model equation. The same study extended the research work further by incorporating viscoelastic material properties to the seat cushion using Prony series formulation.

Exploring the past research works on the automotive seat foam material, it is undoubtedly clear that the polyurethane material can be modelled using the formulations of stress-strain behaviour, hyper-elastic foam properties or combination of hyper-elastic and visco-elastic parameters, though the selection of the formulation in analysis study exclusively depends on the nature of the investigation. Majority of the past research works considered only the hyper-elastic properties for foam materials studies, visco-elastic properties for the assessment of foam material under shear loading or stress-strain curve for monitoring the cushion behaviour. Moreover, those researches considered only certain portion of entire automotive seat or very specific interaction between seat and human body. Very few studies had been found to have considerations for combination of different formulations to model the entire seat, however those studies were conducted using multi-degree of freedom system, lumped mass parameters method or some non-finite element tool. None of the previous case studies had shown the simulation modelling of whole car seat in finite element environment implementing the combined material formulations. Hence, there is a research gap in developing the finite element simulation of entire car seat using polyurethane foam with combined material formulations and therefore, demand arises for a comprehensive simulation based solution to judge the vibration level at any point of entire car seat. Various combinations are possible by taking into account the potential formulations, though careful judgment is necessary to choose the appropriate mathematical model for running flawless computerized simulation. This research paper aims to fill up the gap in existing literatures by outlin-

ing a distinct finite element based simulation set up for entire car seat with the most suitable combination of mathematical formulations for seat foam material.

During this research study, a non-robust car seat has been constructed. The dimensions of the seat are based on the recommended optimized seat parameters and industrial guidelines. During this span of simulation work, combination of hyperelastic and viscoelastic material properties have been taken into account. A comprehensive simulation technique including all the necessary input parameters has been described to establish a simulation model of the entire car seat to assess the vertical vibrations in terms of accelerations at headrest, backrest and cushion.

2. Methodology

2.1. Dimensions of Car Seat

Many research works had been conducted in the past to explore the ideal car seat dimensions to optimize comfort parameters for humans. Fit parameter analysis [15] advised that the least width of cushion to be 432 mm for a 95th percentile female occupant, though increment in the dimension would be beneficial taking into account the clothing. Similar sort of study [16] suggested the minimum width of cushion to be in the range of 480 mm to 500 mm. Based on the anthropometric dimensions, recommended length of seat had also been recorded in different case studies as 432 mm [17], 440 mm to 550 mm [16] and 330 mm to 470 mm [18]. The backrest width and height had been advised to be at least 360 mm and 550 mm above the H-point, respectively, during the survey of automotive seat design [19]. Database of "Ricaro", an industrial innovator in the sector of automobile seat designing, shows how the seat design had been changed over in the last 40 years.

Consulting with all the recommendations to optimize the seat parameters and assuming a 50th percentile male human body to occupy the seat, a CAD model has been established and the major overall dimensions are shown in **Figure 1**.

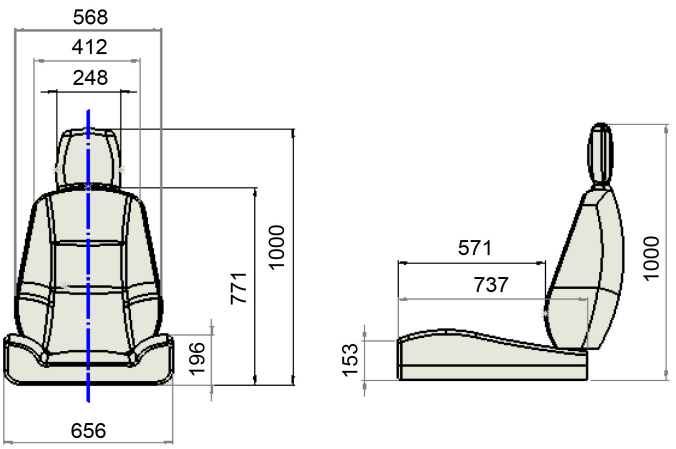


Figure 1. Overall major dimensions of car seat.

2.2. Hyperelastic and Viscoelastic Material Modelling for Seat Foam

Hyperelastic properties are suitable for defining the behaviour of foam material, while the viscoelastic properties are used to judge the effect of shear loading inside the deformable bodies. In this research work, both the hyperelastic and viscoelastic properties have been assigned to the car seat foam cushion to achieve the result as realistic as possible. To find out the best hypothetical formulation to be used, past relevant literatures have been consulted.

In conjunction with international standard ASTM D 3574-01 for testing procedure for flexible foam materials, a relationship between the stress and stress was evaluated [14] by applying second order strain-energy potential function under shear and compression loadings. The hyper-elastic material model was described using the Ogden model formulation [14] as shown in Equation (1).

$$[U]=\sum_{i=1}^{N}2\frac{\mu_i}{\alpha_i^2}\left[\tilde{\lambda}_1^{\alpha_i}+\tilde{\lambda}_2^{\alpha_i}+\tilde{\lambda}_3^{\alpha_i}-3+\frac{1}{\beta_i}\left(\left(J^{el}\right)^{-\alpha_i\beta_i}-1\right)\right] \quad (1)$$

N = Defining parameter for the approximation of the model.

$\tilde{\lambda}_j$ = Principal strength j.

J^{el} = Elastic volumetric ratio.

α_i, β_i, μ_i = Material dependent parameters.

The outcomes of the testing were exported into finite element set up and Ogden hyperelastic coefficients had been estimated by implementing numerical curve fitting technique. The hyperelastic coefficients were obtained as $\mu_1 = 164.861$ kPa, $\alpha_1 = 8.88413$, $\beta_1 = 0$, $\mu_2 = 0.023017$ kPa, $\alpha_2 = 4.81798$, $\beta_2 = 0$. The same investigation further explored the viscoelastic properties of foam and the viscoelastic coefficient values were obtained as $G_1 = 0.3003$, $\tau_1 = 0.010014$ s, $G_2 = 0.1997$, $\tau_2 = 0.10020$ s. Very similar material models were proposed during the performance assessments of slightly compressible hyper-elastic foams [20] and highly compressible hyper-elastic foams [21]. Polynomials and stresses of three kinds of polyurethane foams had been listed [22] to formulate the hyperelastic and viscoelastic material models and later, a series of Ogden coefficients had been tabulated. Finite element analysis of hyperelastic materials also presented [23] a set of Ogden parameters for different scenarios with optimized parameters in numerical and finite element. Polyurethane foam had been modelled using hyper-elastic, visco-elastic, polynomial and stress formulations [24] and a set of data on hyperelastic and viscoelastic coefficients had been reported. Ogden parameter values had been optimized and tabulated as $\mu_1 = 4.81$ kPa, $\alpha_1 = 19.8$, $\beta_1 = 0.01450$, $\mu_2 = 3.60$ kPa, $\alpha_2 = 19.8$, $\beta_2 = 0.0065$ for hyper-elastic materials during the numerical and finite element study [23] of foam material for different case scenarios.

Based on the data gathered from the past works done on the relevant field, hyperelastic Ogden coefficients ($N = 2$) have been implemented to the seat foam. Viscoelastic parameters have been allocated based on the time dependent function of instantaneous shear modulus. The coefficient values used during this re-

search task, are outlined in Table 1 and Table 2.

2.3. Density and Poisson's Ratio

Density of polyurethane foam had been inspected by National Bureau of Standards, USA [25] at low temperatures 295 K, 111 K, 76 K and 4 K and reported as 64 kg/m^3. Car seat cushion modelling using polyurethane foam [1] considered the density value as 67 kg/m^3, while assessment of polyurethane foam material with respect to different types and cell sizes [26] reported the densities in-between 57 kg/m^3 to 68 kg/m^3. While finding the Poisson's ratio for seat cushion foam materials [27], the density had been in the range of 32 kg/m^3 and 64 kg/m^3. Based on the past works carried out on the density values of the foam material, density value of the polyurethane foam for this current analysis task has been taken as 64 kg/m^3.

Latest trend in automotive seat foam design is to develop seat material which will bulge to inward direction and display negative Poisson's ratio. One of the cutting edge technique developed on car seat cushion [27] found the Poisson's ratio of −0.13 for polymers with density of 18 kg/m^3 and compression ratio of 2.2, while with density of 25 kg/m^3 and compression ratio of 3.4, the obtained Poisson's ratio was −0.26. The seat foam material usually under the effects of compression without any sidewise constraints, hence, there is no co-relation between the lateral and longitudinal strains. Taking into account this fact, computational analysis of car seat and human body [14], finite element modelling of the car seat [1] and the process of designing the car seat [13] ignored the Poisson's ratio by assuming its value to be zero. Considering the real fact as described in the revenant investigations, in this simulation project the value of Poisson's ratio has been ignored.

Table 1. Hyperelastic parameters used.

Hyperelastic parameter	Value
μ_1	0.00481 MPa
α_1	19.8
β_1	0.01450
μ_2	0.00360 MPa
α_2	19.8
β_2	0.00650

Table 2. Viscoelastic parameters used.

Viscoelastic parameter	Value
G_1	0.3003
τ_1	0.010014 s
G_2	0.1997
τ_2	0.10020 s

2.4. Simulation Set Up-Loading, Step, Boundary Condition and Meshing

Loading condition in this simulation work has been applied considering a non-racer type accelerating car achieving a speed of 30 miles/hour from standstill condition. The road terrain has been assumed to be smooth, hence, the load primarily has been accounted due to initial accelerating period. Databases of Jaguar XK coupe [28] showed the measured acceleration value as 4.5 m/sec^2. Initial acceleration for the non-racer type of car at 60 km/hour and 40 km/hour [29] reported the acceleration values as 1.083 m/sec^2 and 0.861 m/sec^2, respectively. Average and peak accelerations for a normal passenger vehicle operating in rural area had been formulated [30] through Equation (2) and Equation (3).

$$a_{av} = ae^{bv} \tag{2}$$

$$a_{\max} = c + dv \tag{3}$$

a_{av} = Average acceleration in m/sec^2

$a_{\max}$ = Maximum acceleration in m/sec^2

v = Vehicle speed in m/sec

a, b, c, d = Constants

The car seat in this project work has been assumed to cope with a 50th percentile human male body of 77.3 kg mass. Considering the closest possible matching criteria for vehicle operating scenario, the force exerted on the seat due to acceleration has been calculated using the acceleration value of 0.861 m/sec^2. Contact interfaces between the seat and humans have been estimated and vertical load due to human object and horizontal load due to acceleration have been implemented on the estimated interaction areas, in vertical and horizontal directions, respectively.

The entire simulation set up has been carried out in ABAQUS CAE 6.13. The base condition of the seat has been assigned in "Initial" step, while the loading conditions have been implanted in "Static General" step.

The boundary conditions have been simplified based on the aspects of this project work and implemented as realistic as possible. Underneath the seat cushion has been made fixed, while the backsides of backrest and headrest are not permitted to move from initial state. Angular fore-aft movements of the backrest and headrest about the respective bottom connecting points are allowed.

The common elements used for deformable bodies in finite element are quadrahedral and hexahedral. During this simulation set up, the seat structure has been meshed with ten-node tetrahedral element—C3D10. **Figure 2** is visually representing the anticipated contact surfaces, loading, boundary conditions and meshing.

3. Results

A 64 bit standard computer with Windows XP operating system, RAM of 6 GB and two dual-core 2.1 GHz Intel(R) Pentium(R) CPU B950 had been utilized to

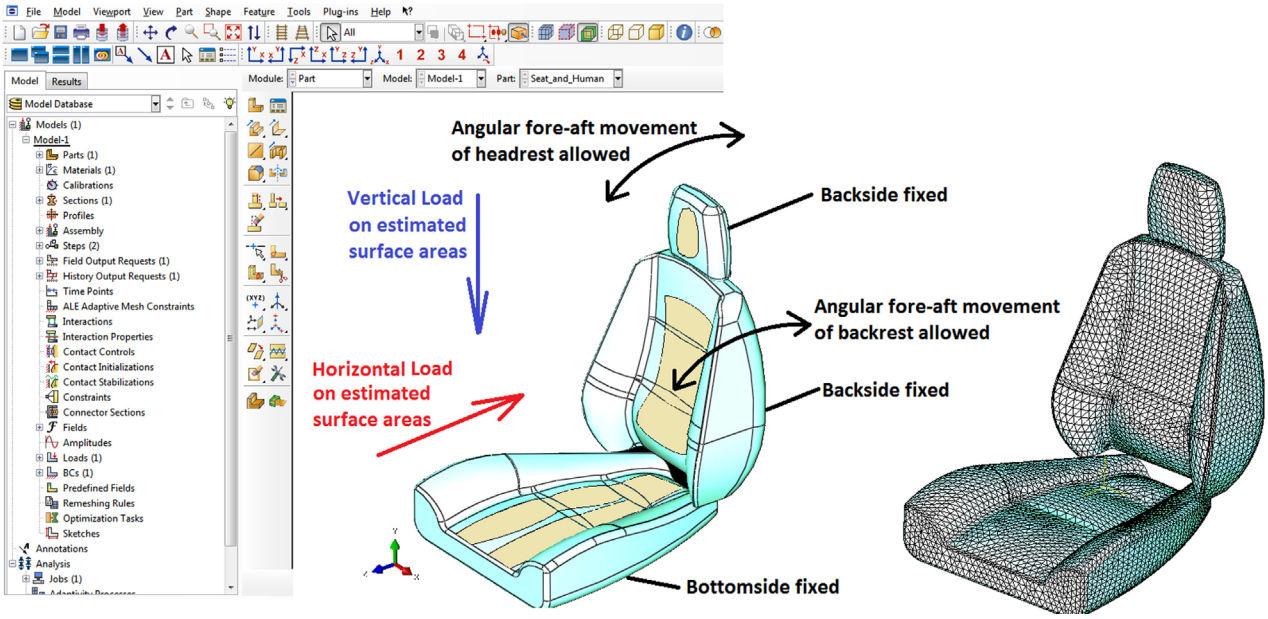


Figure 2. Estimated interfaces, directions of loads, boundary conditions and meshing.

carry out the simulation work. Because of the limitation in computer memory, the simulation running time had been set to 10 seconds. It took around 7 wall clock hours for ABAQUS solver to yield the results. The overall results obtained for acceleration responses, are visually represented in Figure 3.

From the ABAQUS post processor, the vertical accelerations at seat headrest, backrest and cushion had been extracted and shown in Figures 4-6.

The average and RMS acceleration magnitudes of car seat portions have been calculated and shown in Table 3.

4. Validation and Discussion

The vertical accelerations received from this analysis have been validated by comparing to real life test data gathered from identical operating condition to this simulation set up. An economic hatchback car with a male driver of 78 kg sitting on the driver seat had been accelerated from static condition to pick up the speed of 30 - 35 miles/hour and sensors had been mounted on the outer surfaces of headrest, backrest and cushion. There was no sign of irregularities on the road terrain and testing data were logged for 60 seconds. Vibration measuring unit NI 9234 USB module with Compact DAQ chassis and transducer Dytran 3055 were employed to read the acceleration vs time plots.

Transducer was mounted approximately at the central location of designated surface with the help of adhesive tapes and the cable from the transducer was linked to the vibration measuring module NI 9234. A standard laptop installed with signal processing tool "m + p Analyzer" was connected to other side of NI 9234. The test set up is shown in Figure 7.

The testing data were received in .SOT format and through the "m + p Analyzer", the graphs for acceleration with respect to time had been generated as

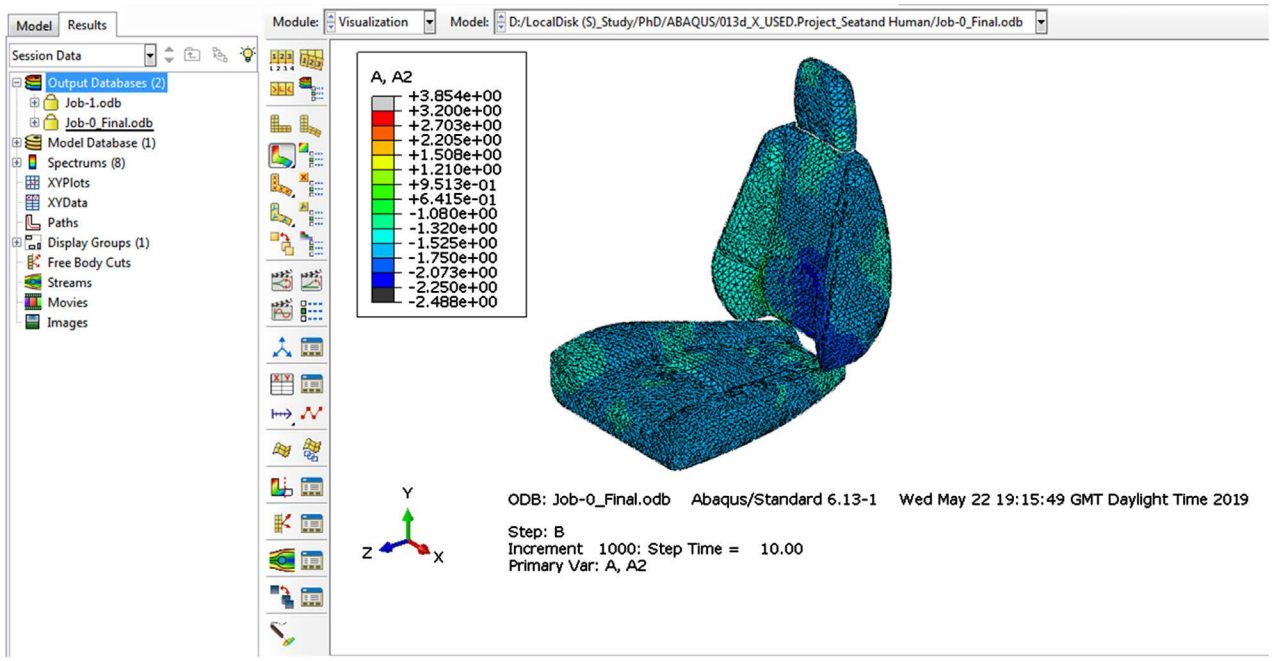


Figure 3. Overall vertical acceleration.

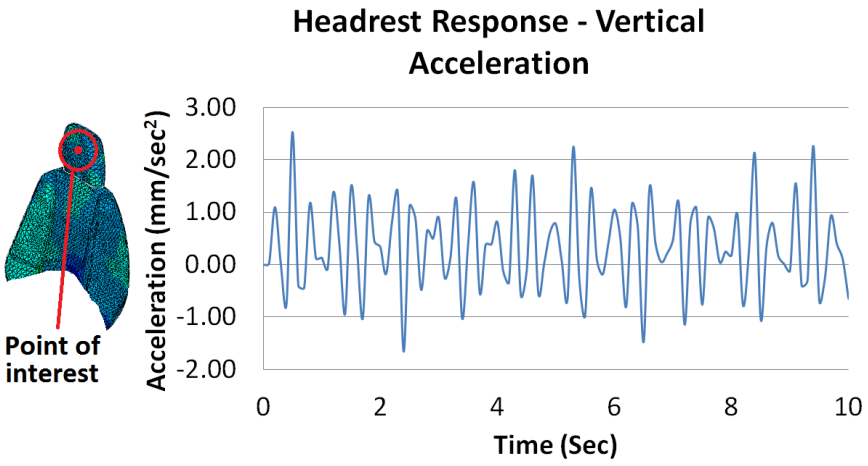


Figure 4. Vertical acceleration at headrest.

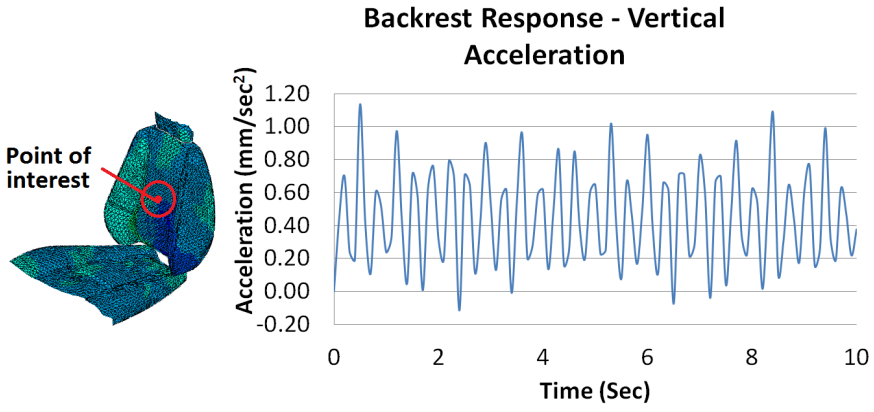


Figure 5. Vertical acceleration at backrest.

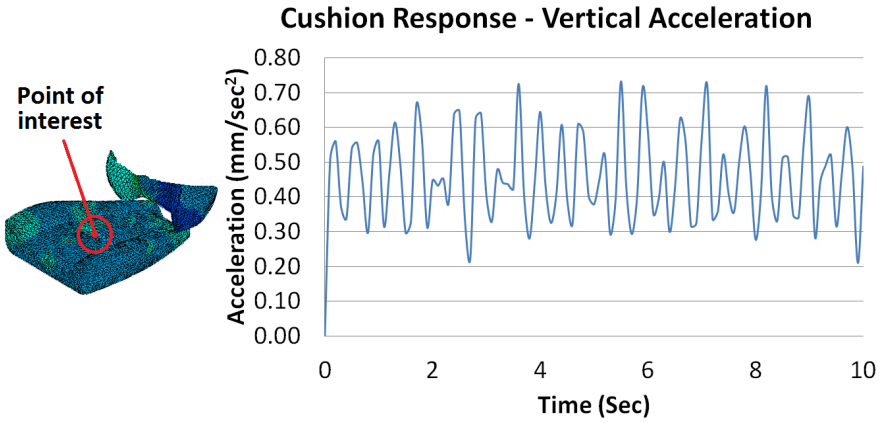


Figure 6. Vertical acceleration at cushion.

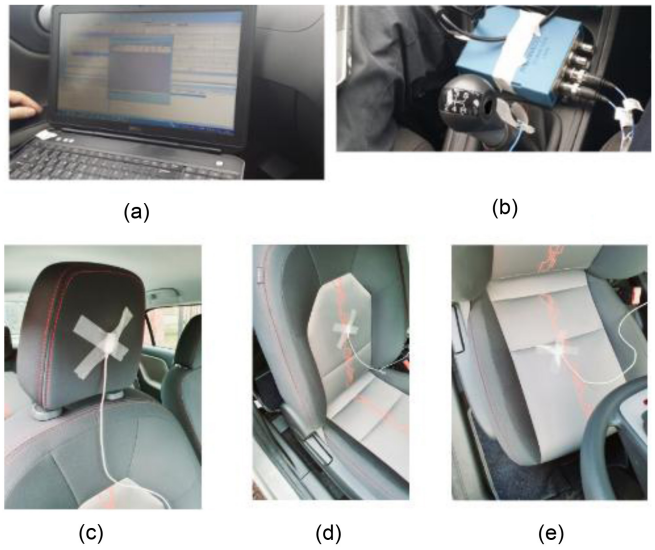


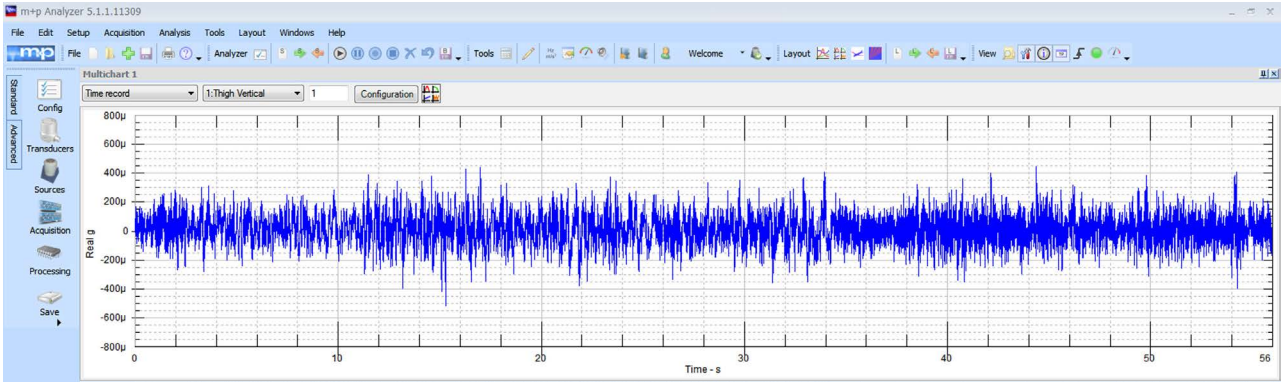
Figure 7. Test set up (a) Laptop with the signal processing tool "m + p Analyzer"; (b) NI 9234 module; (c) Dytran 3055 sensor mounted on headrest; (d) Dytran 3055 sensor mounted on backrest; (e) Dytran 3055 sensor mounted on cushion.

Table 3. Average and RMS accelerations of car seat segments from simulation.

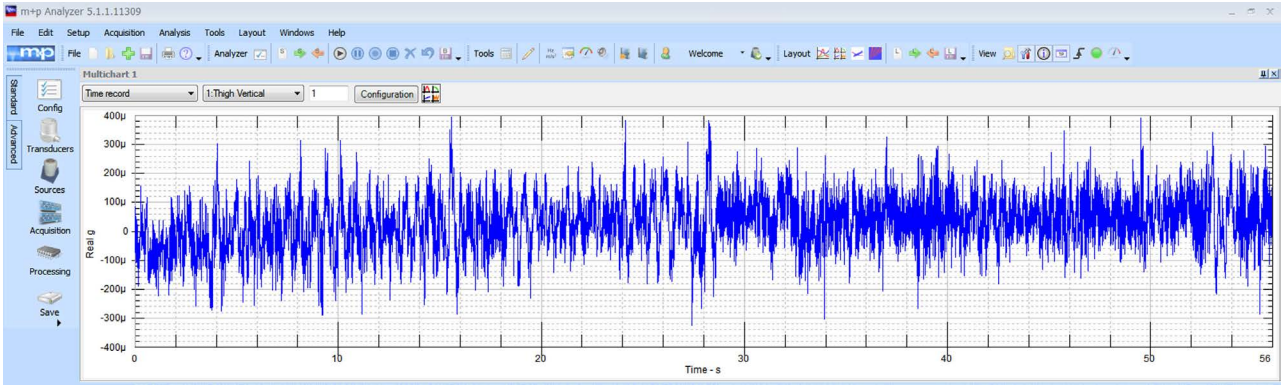
Seat segment	Average acceleration (mm/sec^2)	RMS acceleration (mm/sec^2)
Headrest	0.32	0.91
Backrest	0.45	0.54
Cushion	0.46	0.47

shown in Figure 8.

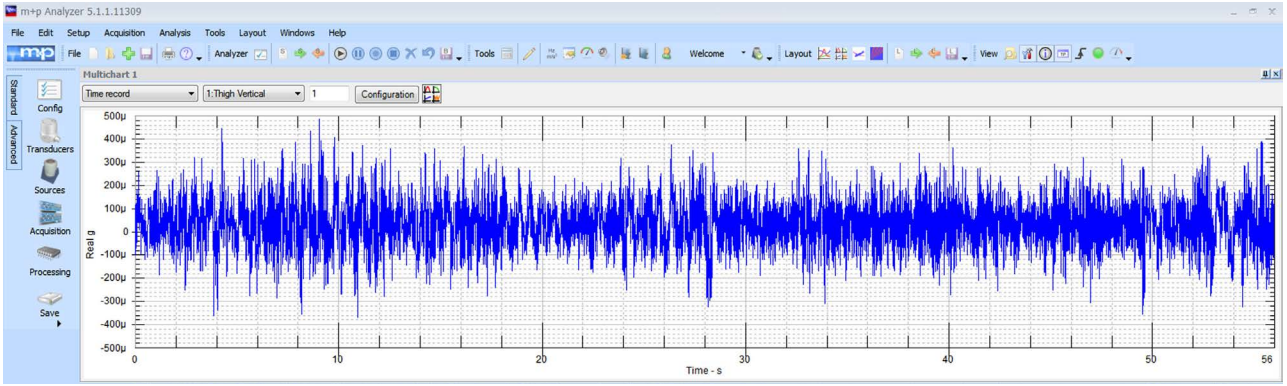
The raw testing data had collected for time period of initial 60 seconds had been curtailed to initial 10 seconds to match the simulation duration and filtered in .XLS format to get clear pictures of the testing data. The average and RMS acceleration magnitudes obtained from test data have been extracted from the graphs and shown in Table 4.



(a)



(b)



(c)

Figure 8. Testing data of vertical acceleration vs time at (a) headrest, (b) backrest, (c) cushion.

Table 4. Average and RMS accelerations of car seat segments from test data.

Seat segment	Average acceleration (mm/sec^2)	RMS acceleration (mm/sec^2)
Headrest	0.02	0.17
Backrest	−0.03	0.16
Cushion	0.04	0.22

Later both the simulation results and testing data were merged into single graphical plot areas for comparison purpose and represented in Figures 9-11.

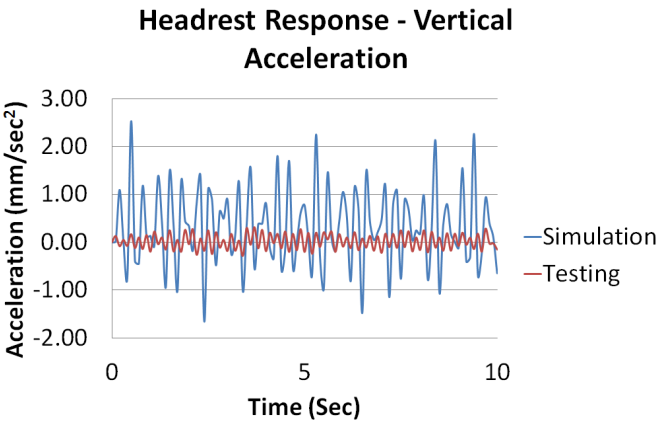


Figure 9. Vertical accelerations of headrest from simulation and testing.

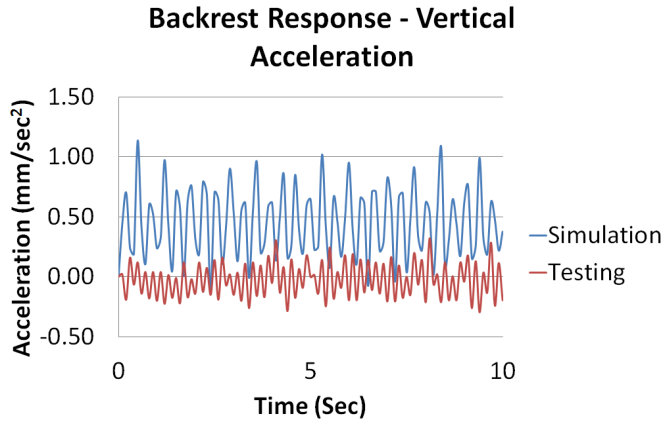


Figure 10. Vertical accelerations of backrest from simulation and testing.

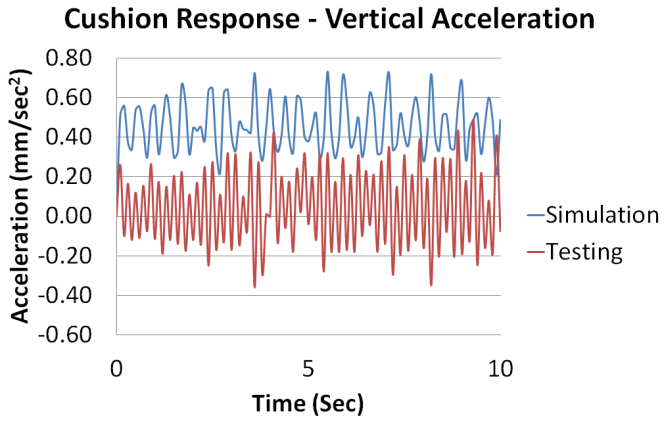


Figure 11. Vertical accelerations of cushion from simulation and testing.

Figure 9 is showing the accelerations of the seat cushion, where both the upper and lower peaks of testing data are well below the respective peaks of simulation results. Figure 10 and Figure 11 are showing the accelerations of headrest

and backrest, respectively, where upper limits of simulation results are exceeding the respective limits of testing data and lower limits of simulation results are well below those of testing data. If the absolute accelerations are considered irrespective of the positive and negative magnitudes of the acceleration values, almost all the simulation results are higher than the testing data, hence, the outputs received from computerized simulation are more conservative in nature.

Deformation mechanism of polyurethane foam structure is a very complex phenomenon where numerous factors are associated and each one of these factors is capable of manipulating the behaviour of foam structure on its own. Type of seat, configuration of seating arrangement, car class, road terrain, time span of acceleration, foam material properties etc. all can be defining parameters for maneuvering the final output from simulation and testing. This fact confirms that the deviations between the simulation results and testing data are inevitable. Moreover, the testing data had been gathered from a car being driven by a real human, while this piece of simulation work considered the loadings coming from virtual human object.

The variations similar to this project task, had been observed in many past research works for automobile and seat related investigations. Vibration transmissions and accelerations were found to be varied with respect to different sitting postures [31] [32] [33] while studies on vibration transmission [34] [35] showed a small alteration in the posture could greatly amend the transfer function and frequency response of the car seated arrangement. Same human body had been used for a number of times in identical test set up [36] and each time the vibration related outputs were different.

The vertical acceleration in the simulation can further be lowered by applying potential remedies as outlined in past relevant literatures on vibration transmission [34] [35] [36] [37]. The vertical acceleration values received from the finite element simulation of this project work are of higher magnitudes than the test data, hence, this simulation set up is useful to optimize the health, comfort and safety levels in worst case scenario. The simulation using both the hyperelastic and viscoelastic materials is able to provide the flawless desired output.

5. Conclusions and Scopes of Further Development

In this research paper, a finite element simulation methodology has been offered to anticipate the final level of vibration by monitoring vertical accelerations at different locations of car seat. From the results of this analysis, the following conclusions can be drawn:

1) This finite element simulation method can be implemented efficiently to the car seat structure to predict the final levels of acceleration responses at different segments. In this simulation set up of car seat, load conditions due to acceleration and human mass along with hyperelastic and viscoelastic material properties have been incorporated. In contrast, in the real life measurement process numerous operating parameters are associated and gathered experimen-

tal data are hugely dependent on the instrument used and its signal processing systems. The basic philosophy behind the vibration measurement system is based on the signal transfer function defined by Laplace domain, where a small alteration in any of the input parameters can tune the final level of vibration to a great extent. So, the mismatches between the test data and simulation output are evident. As the simulation results are more conservative in nature, it can be stated that this simulation technique on car seat using hyperelastic and viscoelastic materials is successful to anticipate the vertical vibrational effect on car seat by means of monitoring acceleration, regardless of the reasonable amount of mismatches between the simulation outputs and testing data.

2) Car seat polyurethane foam can be formulated in computerized simulation by the combination of hyperelastic and viscoelastic materials to run an effective analysis. Assignment of Ogden coefficients for hyperelastic materials and time dependent shear modulus coefficients for viscoelastic materials can lead the simulation to yield necessary outputs.

3) This simulation process is able to predict the acceleration responses inside car seat portions in a sensible way. More advanced research works on the foam material properties are necessary to lower the peak values of accelerations. Numerous combinations of different parameters can be chosen to conduct the simulation and the accurateness of the output will be enhanced with the increment in the number of allocated input parameters.

Enormous potential for further improvement is there to take this emerging simulation methodology to advanced stage. Firstly, a longer finite element simulation running time will inevitably help to obtain enhanced acceleration and frequency plots. Cause of the limitation in computer capability, simulation running time was restricted to 10 seconds during this course of analysis work. Highly configured computer hardware will definitely be beneficial to perform analysis faster and for a longer period of time. Observation on the frequency response plots from this simulation work found that the frequency values are gradually getting stabilized and lowered over the time, hence, longer simulation operating time will eventually display lower ranges of frequencies for the driver human and car seat portions. In case of association of road terrain, non-racer type of automotive and accidental case scenario, it is highly recommended to run the simulation at least for 60 seconds. Secondly, stiffness data can be calculated in a more precise way by subdividing the human segments into smaller portions and viscoelastic coefficients for the time dependent shear modulus function can be introduced along with hyper-elastic material. The frequency response curves received from this simulation are getting steady over time, thus showing the effect of stiffness, damping and hyperelastic material parameters as anticipated. More convincing frequency curves with lower magnitudes can be achieved through improved stiffness data and associated viscoelastic material properties. Thirdly, in the present analysis scenario, only the vertical vibration has been assessed at different segments of the car-driver assembly. For the completeness in understanding of the vibration related effects, it will be beneficial to

take account of the fore-aft vibration by means of acceleration and displacement along the direction of vehicle movement. Comprehensive solution based on cost effective simulation process for predicting accelerations and frequencies at different points of human-car system can omit the necessity of time consuming real life vibration testing procedure.

Based on all the discussions on simulation results and validation, it can be concluded that this unique biodynamic simulation methodology in finite element environment is presenting realistic output data regardless of the fact that the magnitudes of the peak responses are out of the permissible criteria.

Therefore, this simulation technique presented in this research paper is following the right path and recommended fine-tunings on different parameters can successfully lead this technology to anticipate accurate acceleration and frequency responses of full car seat and human system.

Acknowledgements

Simulation results have been validated with the practical testing data provided by "m + p International". Valuable suggestions and recommendations are further given by "m + p International" on the type of outputs required from the simulation for effective validation of finite element results. Authors of this research paper are indebted to "m + p International" for all their help and supports.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Haan, R. (2002) FE Model of a Car Seat. Netherlands Organisation for Applied Scientific Research, 9-12.

[2] Singh, R., Davies, P. and Bajaj, A.K. (2003) Identification of Nonlinear and Viscoelastic Properties of Flexible Polyurethane Foam. *Nonlinear Dynamics*, **34**, 319-346. https://doi.org/10.1023/B:NODY.0000013511.07097.87

[3] Dorugade, D.V., Rakheja, S. and Boileau, P.E. (2019) Modeling and Validation of Static and Dynamic Seat Cushion Characteristics. 12*th European LS-DYNA Conference*, Koblenz, Germany, 14-16 May 2019.

[4] Zhao, L.Q., Xia, Q.S. and Wu, X.T. (1994) Study of Sitting Comfort of Automotive Seats. *SAEConference* 1994, SAE No. 945243.

[5] Park, S. and Kim, C. (1997) The Evaluation of Seating Comfort by Objective Measurements. SAE970595. https://doi.org/10.4271/970595

[6] Xu, W., Zeng, Y. and Ye, J. (2019) Study on Virtual Simulation Method of Driver Seat Comfort. *Proceedings of* 2*nd International Conference on Frontiers of Materials Synthesis and Processing*, **493**, 1-6. https://doi.org/10.1088/1757-899X/493/1/012096

[7] Warner, C.Y., Stother, C.E., James, M.B. and Decker, R.L. (1991) Occupant Protection Inrear-End Collisions: II. The Role of Seat Back Deformation in Injury Reduc-

tion. *Proceedings of the* 35*th Stapp Car Crash Conference* 1991, San Diego, CA, 18-20 November 1991, 379-390.

[8] Qiu, D., He, Y. and Yu, Z. (2019) Investigation on Compression Mechanical Properties of Rigid Polyurethane Foam Treated under Random Vibration Condition: An Experimental and Numerical Simulation Study. *Materials*, **12**, 1-17. https://doi.org/10.3390/ma12203385

[9] Rusch, K.C. (1965) Dynamic Behaviour of Flexible Open-Cell Foams. Ph.D. Thesis, University of Akron, Akron, OH.

[10] Zhang, J., Kikuchi, N., Li, V., Yee, A. and Nusholtz, G. (1998) Constitutive Modeling of Polymeric Foam Material Subjected to Dynamic Crash Loading. *International Journal of Impact Engineering*, **21**, 734-743. https://doi.org/10.1016/S0734-743X(97)00087-0

[11] Choi, H.Y., Lee, W.R., Park, J.C. and Yang, K.Y. (2018) Riding Comfort Simulation with Air Ride Seat for Heavy Duty Vehicle. *Proceedings of the* 2*nd Japanese Modelica Conference*, Tokyo, Japan, 17-18 May 2018.

[12] Dahil, L., Karabulut, A., Baspinar, M.S. and Mutlu, I. (2016) Investigation of Vibration Damping in the Passenger Seat. *The Online Journal of Science and Technology*, **6**, 52-57. https://doi.org/10.17932/IAU.IJEMME.m.21460604.2016.5/1.1117-1122

[13] Camprubí, N. and Rueda, F. (2007) Comfort Evaluation of Foam Seats Using Realistic Simulation. Advanced Design & Analysis Division.

[14] Grujicic, M., Bell, W.C., Arakere, G. and Haque, I. (2009) Finite Element Analysis of the Effect of Up-Armouring on the off-Road Braking and Sharp-Turn Performance of a High-Mobility Multi-Purpose Wheeled Vehicle. *Proceedings of the Institution of Mechanical Engineers Part D-Journal of Automobile Engineering*, **223**, 1419-1434. https://doi.org/10.1243/09544070JAUTO1187

[15] Gordon, C.C., Churchill, T., Clauser, C.E., Bradtrniller, B., McConville, J.T., Tebbetts, I. and Walker, R.A. (1989)1988 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics. Final Report (NATICWR-891027), U.S. Army Natick Research, Development and Engineering Center, Natick, MA.

[16] Grandjean, E. (1980) Sitting Posture of Car Drivers from the Point of View of Ergonomics. In: Oborne, D.J. and Levis, J.A., Eds., *Human Factors in Transport Research. User Factors*: *Comfort*, *the Environment and Behaviour*, Academic Press, New York, 205-213.

[17] Keegan, J.J. (1964) The Medical Problem of Lumbar Spine Flattening in Automobile Seats. SAE Technical Paper 838A. Society of Automotive Engineers, Inc., New York. https://doi.org/10.4271/640788

[18] Chaffm, D.B. and Anderson, G.B. (1991) Occupational Biomechanics. 2nd Edition, Wiley-Interscience, New York.

[19] Reed, M.P., Schneider, L.W. and Ricci, L.L. (1994) Survey of Auto Seat Design Recommendations for Improved Comfort. UMTRI, 5-10.

[20] Ogden, R.W. (1972) Large Deformation Isotropic Elasticity-On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids. *Proceedings of the Royal Society of London Series A*, **326**, 565-584. https://doi.org/10.1098/rspa.1972.0026

[21] Mills, N.J. (2007) Polymer Foams Handbook: Engineering and Biomechanics Applications and Design Guide. Butterworth-Heinemann, Waltham, MA.

[22] Ju, M.L., Jmal, H., Dupuis, R. and Aubry, E. (2014) Visco-Hyperelastic Constitutive Model for Modelling the Quasi-Static Behavior of Polyurethane Foam in Large De-

formation. *Polymer Engineering and Science*, **55**, 1795-1804. https://doi.org/10.1002/pen.24018

[23] Schrodt, M., Benderoth, G., Kuhhorn, A. and Silber. G. (2005) Hyperelastic Description of Polymer Soft Foams at Finite Deformations. *Technische Mechanik*, **25**, 162-173.

[24] Ju, M.L., Jmal, H., Dupuis, R. and Aubry, E. (2013) Visco-Hyperelastic Model for Polyurethane Foam: Comparison among Polynomial, Reduced Polynomial, and Ogden Models. In: 21*ème Congrès Français de Mécanique*, Laboratoire MIPS, Mulhouse, France, 26 au 30 août. https://doi.org/10.4028/www.scientific.net/AMR.856.169

[25] Arvidson, J.M., Sparks, L.L. and Guobang, C. (1983) National Bureau of Standards. Tensile, Compressive and Shear Properties of a 65-kg/m3 Polyurethane Foam at Low Temperatures. https://doi.org/10.6028/NBS.IR.83-1684

[26] Jarfelt, U. and Ramnäs, O. (2006) Thermal Conductivity of Polyurethane Foam Best Performance. *Proceedings of the* 10*th International Symposium on District Heating and Cooling*, Hannover, Germany, 3-5 September 2006, 1-12.

[27] Lakes, R.S. and Lowe, A. (2000) Negative Poisson's Ratio Foam as Seat Cushion Material. *Cellular Polymers*, **19**, 157-167.

[28] Wardell, G. (2007) Jaguar XK Coupe Review. The Auto Channel.

[29] Mehar, A., Chandra, S. and Velmurugan, S. (2013) Speed and Acceleration Characteristics of Different Types of Vehicles on Multi-Lane Highways. *European Transport\Trasporti Europei*, **55**, 1-12.

[30] Brooks, R.M. (2012) Acceleration Characteristics of Vehicles in Rural Pennsylvania. *International Journal of Recent Research and Applied Studies*, **12**, 449-453.

[31] Hinz, B. and Seidel, H. (1987) The Non-Linearity of the Human Body's Dynamic Response during Sinusoidal Whole Body Vibration. *Industrial Health*, **25**, 169-181. https://doi.org/10.2486/indhealth.25.169

[32] Panjabi, M.M., Andersson, G.B.J., Jorneus, L., Hult, E. and Mattsson, L. (1986) *In Vivo* Measurement of Spinal Column Vibrations. *Journal of Bone and Joint Surgery*, **68**, 695-702. https://doi.org/10.2106/00004623-198668050-00009

[33] Mansfield, N.J. and Griffin, M.J. (2000) Non-Linearity in Apparent Mass and Transmissibility during Exposure to Whole-Body Vertical Vibration. *Journal of Biomechanics*, **33**, 933-941. https://doi.org/10.1016/S0021-9290(00)00052-X

[34] Kitazaki, S. and Griffin, M.J. (1998) Resonance Behaviour of the Seated Human Body and Effects of Posture. *Journal of Biomechanics*, **31**, 143-149. https://doi.org/10.1016/S0021-9290(97)00126-7

[35] Zimmermann, C.L. and Cook, T.M. (1997) Effects of Vibration Frequency and Postural Changes on Human Responses to Seated Whole-Body Vibration Exposure. *International Archives of Occupational and Environmental Health*, **69**, 165-179. https://doi.org/10.1007/s004200050133

[36] Griffin, M.J. (1990) Handbook of Human Vibration. Academic Press, London.

[37] Pope, M.H., Broman, H. and Hanson, T. (1990) Factors Affecting the Dynamic Response of the Seated Subject. *Journal of Spinal Disorders*, **3**, 135-142. https://doi.org/10.1097/00002517-199006000-00004

Scientific Research Publishing

Engineering, 2020, 12, 194-215
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

Tourism Traffic Demand Prediction Using Google Trends Based on EEMD-DBN

Yi Xiao[1*], Xueting Tian[1], Ming Xiao[2]

[1]School of Information Management, Central China Normal University, Wuhan, China
[2]Network Center, Central China Normal University, Wuhan, China
Email: *yxiao@mail.ccnu.edu.cn

How to cite this paper: Xiao, Y., Tian, X.T. and Xiao, M. (2020) Tourism Traffic Demand Prediction Using Google Trends Based on EEMD-DBN. *Engineering*, **12**, 194-215.
https://doi.org/10.4236/eng.2020.123016

Received: January 29, 2020
Accepted: March 27, 2020
Published: March 30, 2020

Abstract

Predicting tourism traffic demand accurately plays an important role in making effective policies for tourist administration. It helps to distribute the resources reasonably and avoid the tourism congestions. This paper considered the noise interference and proposed a hybrid model, combining ensemble empirical mode decomposition (EEMD), deep belief network (DBN) and Google trends, for tourism traffic demand prediction. This model firstly applied dislocation weighted synthesis method to combine Google trends into a search composite index, and then it denoised the series with EEMD. EEMD extracted the high frequency noise from the original series. The low frequency series of search composite index would be used to forecast the low frequency tourism traffic series. Taking the inbound tourism in Shanghai as an example, this paper trained the model and predicted the next 12 months tourism arrivals. The conclusion demonstrated that the forecast error of EEMD-DBN model is lower remarkably than the baselines of ARIMA, GM(1,1), FTS, SVM, CES and DBN model. This revealed that nosing processing is necessary and EEMD-DBN forecast model can improve the prediction accuracy.

Keywords

Tourism Traffic Demand Forecasting, Deep Learning, Google Trends, Composite Search Index, Ensemble Empirical Mode Decomposition (EEMD), Deep Belief Network (DBN)

1. Introduction

According to data released by China Tourism Research Institute, the growth rate of inbound tourist volumes in China is relatively slow. That is to say, for a long

time in the past, the development of China's inbound tourism has been basically stagnant, which is inconsistent with the hot situation of the domestic tourism market and outbound tourism. The tourism demand forecast can provide timely basis for relevant departments to formulate effective tourism policies [1]. Modern information technology has brought great convenience to people's life, work and study. People usually turn to search engines when making travel strategies, and people's travel plans often use keyword search information as a reference [2].

In recent years, the research results of using the network search data to establish the tourism demand forecasting model are quite fruitful. The traditional econometric model and the machine learning method will be limited by historical data when forecasting the tourism demand, compared with the network search data. Search has instantaneity and subjectivity and more accurately the needs of tourists can be reflected. In fact, as early as 2009, the predictive power of web search data has been confirmed, for example, the application of Google Trends in all walks of life is considered effective [3].

Tourism is a related industry and has been greatly affected by emergencies; it has been difficult to solve the impact of emergencies on the tourism industry. The forecasting model cannot be adjusted in real time according to the changes in tourism dynamics, for example, the congestion of tourists at the entrance of Jiu Zhaigou Scenic Spot in Sichuan, China, on October 2, 2013 and Shanghai Bund stampede, on December 31, 2014 etc. These incidents have led that the number of tourists in local tourist attractions is difficult to dredge, and the quality of service of scenic spots has declined. It has become a hot spot of concern to all sectors of society. These problems indicate that the spatial allocation of tourism resources is crucial to the healthy development of tourism [1]. It is possible to balance tourism volumes of different tourist attractions due to the instant and efficient tourism demand forecasting model. In this way, the tourism industry will be more orderly and standardized, creating a pleasant atmosphere for China's inbound tourism.

In the field of tourism demand forecasting, research methods will vary depending on the conditions and objects of the forecast. Inbound tourists are more purposeful. Due to long distance of travel and the relatively long stay time compared to domestic tourism, the possibility of planning ahead is even greater. It is more prevalent to rely on online search to develop a travel schedule. However, the prediction model established by simply using the search data is not robust, and the combination of artificial intelligence and search data can greatly improve the accuracy of prediction [4] [5]. This paper uses EEMD to decompose the historical tourist volumes sequence of Shanghai inbound tourism and Google keyword search data respectively to eliminate the adverse effects of noise interference on the prediction results. Finally, the DBN with better convergence effect is used to predict the tourist volumes with the synthetic search index, which ensures the real-time validity of the prediction model.

2. Literature Review

2.1. Tourism Demand Forecast

In the forecast of tourism demand, there have been very rich research results in the past ten years [6]. Whether different methods or combinations have better predictive effects have always been the main direction of scholars' exploration [7]. The methods currently used in this field can be roughly divided into traditional time series models, artificial intelligence prediction methods [8] [9] and hybrid methods [10]. Among the studies of time series models, the most used ones are ARIMA models [11] [12], exponential smoothing models [13], and linear regression [14] and so on. Among them, ARMA has diverse prediction performance under different conditions, that is, it can be adjusted according to different research conditions to achieve better prediction results. ARMA has more possibilities [15]. The exponential smoothing model has also gradually evolved from primary exponential smoothing to quadratic exponential smoothing to cubic exponential smoothing to obtain more accurate predictions. In the study of tourism demand forecasting in several major source countries in Australia, we can see a comparison of several exponential smoothing models [12]. In fact, it can be observed in the research of many scholars that no one method has an absolute advantage. In general, we think that the combined model is more accurate than the single model [4] [5].

The application of artificial intelligence methods in tourism demand forecasting has begun to rise in the past 30 years. A back propagation neural network model can be applied to tourism demand forecasting [16]. As one of the international tourist cities, there have many methods for the prediction of Hong Kong's demand for inbound and outbound tourism, such as rough set theory [17]. Grey models are also widely used in tourism demand forecasting, including research on air passenger traffic [18] [19]. Genetic algorithms [20] have been developed from artificial neural networks [21]. The two major source countries of the Balearic Islands, the UK and Germany, have corresponding visitors every month, and some scholars have used genetic algorithms to conduct special research on this, which shows that genetic algorithms are also feasible in tourism demand forecasting [22]. Support vector machines can better solve practical problems such as small samples, nonlinear, high-dimensional numbers and local minimum points, adding network search data will greatly improve the accuracy of the prediction model, which predicts the passenger flow of Barbados Island, and there is a good embodiment [23].

2.2. Network Search Forecast

Since the network search data are used to successfully predict the epidemic [24], it has begun to use the search data to predict the phenomenon in many fields such as economics and social sciences. For example, scholars used Google search data to predict unemployment, housing prices, stocks [25] [26] [27], etc. Web search data has indeed contributed its valuable role in research in various fields,

especially in today's rapid development of information technology, when people are increasingly relying on online query tools we also hope to use Google search data to further study future consumer behavior [28].

The application of network search data is more and more extensive, providing a good enlightenment for the research of the tourism industry [28] [29]. However, due to the different language and cultural background of each region or country, the search intensity of multi-language source market sets and different leading search engine platforms will also affect the results of tourism demand forecasting [30]. A nonlinear auto-regressive method is combined with keyword search data to predict Malaysia's passenger volumes shows good predictive performance [28]. Since then, Google search data has also been used in the tourism demand forecasting study in the Caribbean region [31]. It indicates the effectiveness of web search data applications in forecasting tourism demand.

In China, there are numerous users of Baidu search engines [32]. Most of the scholars' research is dependent upon the Baidu search index. However, in the world, Google search engine dominates. This article takes Shanghai's inbound tourism demand forecast as an example, targeting the world's tourist groups, so it uses the data of the Google search engine.

3. Methodology Formulation

3.1. Principle of EEMD (Ensemble Empirical Mode Decomposition)

EEMD is an improved algorithm of EMD (Experimental Mode Decomposition) [33], which effectively solves the problem that EMD relies on local number of extreme data information. EMD is generally used for the decomposition of the original sequence from the data itself, because EMD is decomposed according to its own characteristics, and no other prior conditions are needed, so it is more used in noise processing and prediction. However, when the signal is not stable or contains anomalous events, EMD cannot show its superiority [34]. When the signal is disturbed by an abnormal event (such as pulse interference); mode mixing phenomenon occurs. In order to compensate for the shortcomings of EMD in modal decomposition, EEMD can be effectively solved to solve the model aliasing phenomenon. EEMD can make the decomposition scale more uniform, suppress the influence of abnormal events on the signal, and make the prediction more accurate.

The basic methods of EEMD are as follows:

Step 1: Calculate the sequence (set to $P_{(t)}$) local number of extreme data point using EMD, the maximal value constitutes the upper envelope $m_{(t)}$, and the minimum value constitutes the envelope $n_{(t)}$, and the mean $z_{(t)}$ of the upper and lower envelopes at any point is zero.

Step 2: Subtract the mean of the upper and lower envelopes with the sequence to get $R_{(t)}$.

$$R_{(t)} = \frac{P_{(t)} - \left(m_{(t)} + n_{(t)}\right)}{2} \tag{1}$$

Verify that $R_{(t)}$ satisfies the IMF. If not, repeat steps 1 and 2 until $R_{(t)}$ satisfies the IMF condition, and treat $R_{(t)}$ as an IMF separated from $P_{(t)}$ one by one. In the above process, the finite number of IMF_i components and the sum of the remainders $u_{(t)}$ and $Y_{(t)}$ are decomposed one by one from high frequency to low frequency by multiple screenings.

$$Y_{(t)} = \sum_{i=1}^{N} IMF_i + u_{(t)} \tag{2}$$

Step 3: Add random white noise to the sequence $P_{(t)}$, and equalize the abnormal events, so that the abnormal event mode is mixed into the random white noise mode during the EMD decomposition process, and then normalized, and the random white noise is applied by applying the EMD pair. The subsequent signal is decomposed to obtain an IMF_i component.

Step 4: Get IMF_i integration after decomposition (adding a new random normal distribution white noise sequence)

$$P_{(t)} = \sum_{i,j=1}^{n} IMF_{i_{(t)},j_{(t)}}, i = 1,\cdots,N; j = 1,\cdots,n \tag{3}$$

Step 5: Preset a threshold k. If the integrated value in the fourth step is less than k, it is removed as noise. If the integrated value in the fourth step is greater than k, the IMF_i is reset, and Q is an entropy function.

$$K_i\left[IMF_{i_{(t)}}, P_{(t)}\right] = Q\left[IMF_i\right] - Q\left[IMF_{i_{(t)}}, P_{(t)}\right] \tag{4}$$

3.2. Compositions of Google Search Keyword Variables

The era of big data brings new opportunities for the establishment of tourism demand forecasting models. The dependence of users on search engines can provide important data for tourism demand forecasting. In this paper generalized dynamic factor model (GDMF) [35], which can process high-dimensional data, is used to combine keyword variables. The unique advantage of GDMF is that variable data can be updated in real time, and known variables can be interpreted by common parts of unknown variables, aggregated into travel-related indices [36].

Forni (2004) proposed the idea of using VAR to represent the model of GDFM [37]. The traditional factor model is composed of the sum of s common factor k_t and special factor j_t. On this basis, GDFM gives common factors. Partially multiplied by the load matrix of m^*n, denoted as α, then the observed variable can be expressed as:

$$X - \alpha k_t + j_t \tag{5}$$

The matrix transformation of k_t can be expressed as:

$$k_t = pk_{t-1} + q\delta_t \tag{6}$$

where $\delta_t = (\delta_{1t}, \delta_{2t}, \delta_{3t}, \cdots, \delta_{nt})$ is a s-dimensional common component.

Tourism demand has many uncertainties and is able to be influenced by policies and media indices. The network search can reflect the tourists' decision-making behavior motives, but due to the influence of some emergencies, the search volume of one or several keywords in a certain period will be extremely high or very low, and these data are abnormal. So we create a standard scale for the search data when synthesizing the keywords. When the data show a maximum value beyond the standard scale, the method of taking the mean value is used to process the abnormal data. Using EEMD to decompose tourist volumes sequence n IMF components, the same approach is applicable to the Google keyword search index sequence.

3.3. DBN (Deep Belief Network) Prediction Model

Hinton proposed the Deep belief network [38], and the initial parameters of the model are obtained through unsupervised training methods. Compared with the traditional neural network model, DBN does not need a large number of supervised signals, and is not easy to fall into local minimum, which can greatly improve the convergence efficiency. The DBN network model is stacked by multiple restricted Boltzmann machine (RBM) layers. The DBN model with the best effect is obtained through repeated training on multiple layers RBM. The most basic RBM consists of a hidden layer and a visible layer.

Hinton proposed the idea of training each layer of RBM separately [39], which is, extracting input data features in the hidden layer of the first layer RBM, using it to train the second layer RBM, repeating this process until DBN all RBMs stacked in the model are trained. It is assumed that the established DBN model has a total of c hidden layers and d visible layers, and the values of the i^{th} visible layer and the j^{th} hidden layer are: $<v_i, h_j>$, and the bias of the two is: a_i, b_j, then the parameter θ of the RBM is:

$$\theta = (\omega_{ij}, a_i, b_j) \tag{7}$$

Then, the energy of the RBM is expressed as:

$$E(v, h : \theta) = -\sum_{i=1}^{c} a_i v_i - \sum_{j=1}^{d} b_j h_j - \sum_{i=1}^{c}\sum_{j=1}^{d} a_i \omega_{ij} h_j \tag{8}$$

ω_{ij} is the symmetric connection weight between the visible layer v_i and the hidden layer h_j. The probability of the binary state of the hidden layer v_i being set to 1 and the probability of the binary state of the visible layer h_j being set to 1 are calculated [39]. In general, the algorithm of contrast divergence is used to represent the log likelihood gradient of RBM, and the weight and offset parameters are updated by calculation, for detailed algorithm, see Hinton's work [40].

In this paper, we propose an EEMD-DBN prediction model whose structure is shown in **Figure 1**.

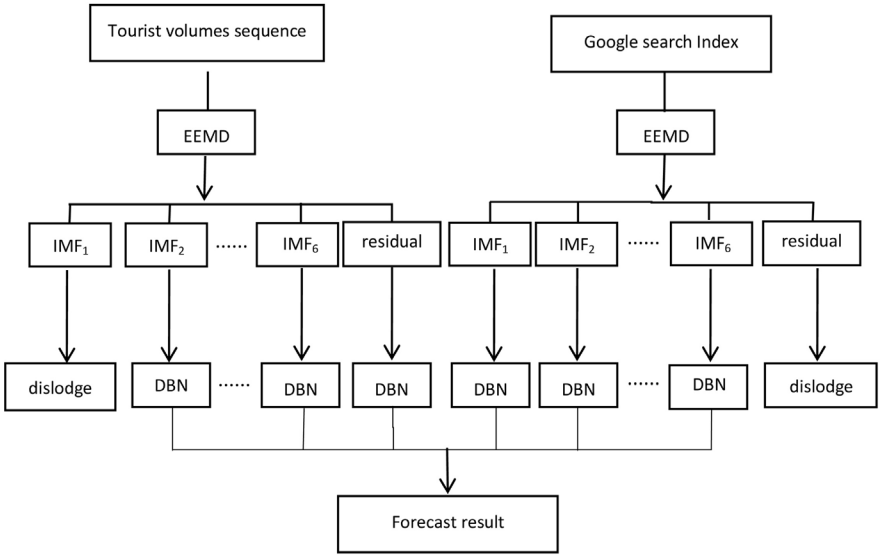


Figure 1. The structure of EEMD-DBN prediction model.

4. Experimental Study

4.1. Data Set

4.1.1. Tourist Volumes Data

Shanghai is an important pillar of China's national economic development; and it attracts many tourists from home and abroad to visit here with rich tourism resources. The regional differences in China's inbound tourism development have gradually narrowed, but it is undeniable that Shanghai still has a huge impetus to the progress of China's inbound tourism. According to the 2017 National Statistical Report on National Economic and Social Development, as of the end of 2017, there were 99 A-level scenic spots in Shanghai, including 3 scenic spots in 5A, 50 scenic spots in 4A, and 46 scenic spots in 3A. It has become one of the cities with the most inbound tourist volumes in China.

The data selected in this paper are the number of monthly inbound tourists from 2004 to 2018 in Shanghai, and divide the data into two parts: the training set and the predictive test set. In order to ensure the validity of the prediction model, 2004-2017 was selected as the sample data for the training and establishment of the prediction model, and the 2018 tourist data were used as the prediction set. Baidu search engine is more widely used in China, so it is more suitable for China's domestic tourism demand forecast. Globally, Google's users are more extensive, accounting for about 66.7% of the world's total [29]. So this article uses the search data of Google search engine. **Figure 2** is a time-series map of monthly inbound tourist traffic in Shanghai from 2004 to 2017. As can be seen from the figure, the time series of tourist traffic shows an overall slow upward trend and cyclical fluctuations in a certain period of time. And in the third quarter and fourth quarter of 2010, there is a peak feature, which is inseparable from the occurrence of the Shanghai World Expo in the second quarter of 2010. The monthly data used in this paper has a total of 168 data points, which better

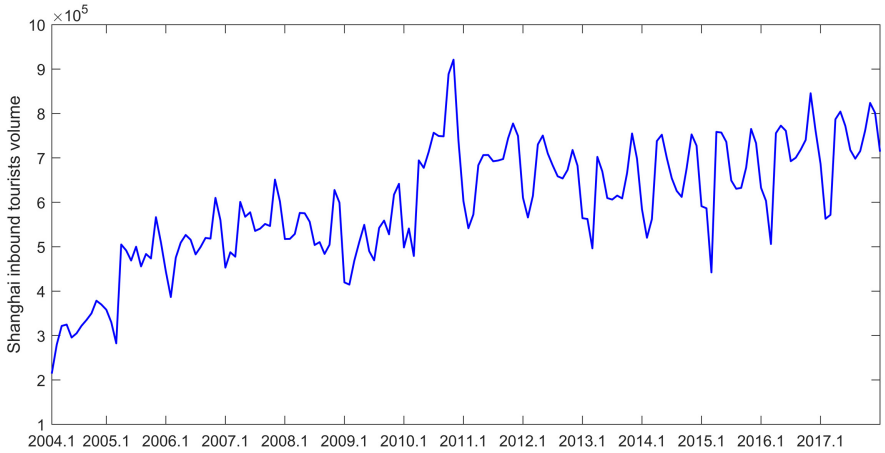


Figure 2. Tourist volumes sequence in Shanghai.

reflects the long-term trend of Shanghai tourist volumes. Compared with the annual data, the time series of monthly data changes more significantly, which can provide a more detailed basis for tourism destination tourism decisions.

4.1.2. Keyword Search Data

Compared with short-distance travel, inbound tourists will stay at the destination for a relatively long time, so people have a tendency to use search engines to develop relevant travel plans in advance. This includes travel route planning, travel hotel ordering, and travel destination information inquiry. This series of behaviors is basically done by means of search engines, so this paper uses keyword search index to continue to improve the forecasting accuracy of tourism demand forecasting model. A key step in the synthesis of online search index is the selection of search keywords. In terms of keyword selection, there is currently no mature program and theoretical system. This article uses a more common method of directly selecting keywords.

Firstly, the 50 common keywords related to tourism are selected for search volume search. According to the Google search volume ranking, the first 16 keywords are retained, and the Pearson correlation test is performed on these 16 keywords, and the correlation with the tourism is the largest 5 keywords as research samples as **Figure 3**. There is some abnormal value in the network search volume of the 5 keywords we finally determined. Directly eliminating the abnormal data will lead to the lack of research samples. Therefore, this paper performs the mean processing on the abnormal data values to ensure the integrity of the experimental data.

4.2. Data Analysis

It can be seen from the time series of inbound tourist traffic in Shanghai that the fluctuation of passenger tourism volumes data is more obvious, and there is a large amount of data to be processed. The peak of the tourism volumes sequence is from September to October of 2010. As we all know, from May to June 2010,

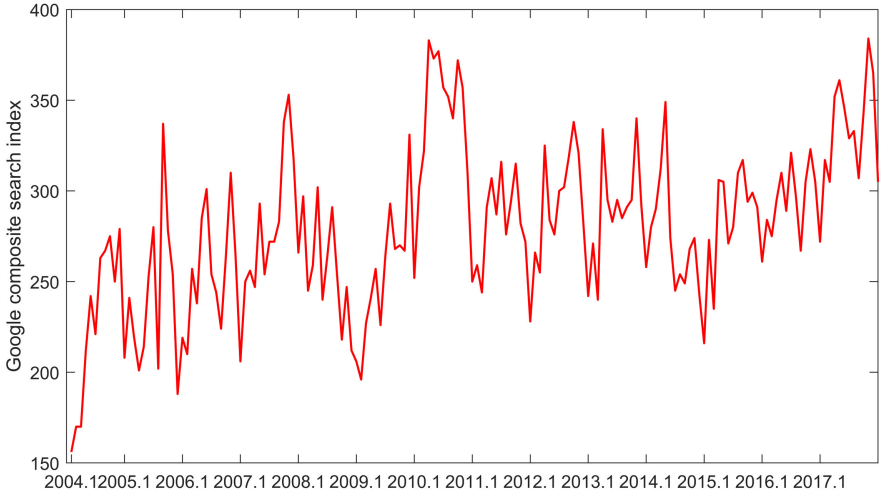


Figure 3. Keyword search data sequence.

Shanghai hosted the World Expo, which played a positive role in Shanghai's tourism development to a certain extent, which led to the rapid growth of Shanghai's inbound tourism volumes in a short period of time. In order to reduce the impact of abnormal events on the forecasting model, we performed a simple noise reduction process on the tourism volumes time series to eliminate the interference of the data peaks and valleys. As shown in **Figure 4**, the time series after smoothing the linear trend factor of the volume data is relatively flat.

Determine keywords according to the eight characteristics of tourism activities, eat, live, travel, travel, purchase, entertainment, determine the basic keywords, such as: travel to Shanghai, Shanghai attractions, Shanghai hotels, Shanghai flights, etc., then enter the basic keywords for Google search volume Inquire. When entering the basic keywords on Google, Google will intelligently recommend keywords related to it, and then record relevant keywords recommended by Google, and sort out 50 keywords related to inbound tourism in Shanghai, the 16 keywords with the largest search volume are retained, and the Pearson correlation coefficient between the tourism volumes and the search keyword is calculated. From the analysis results, we can see that there is a negative correlation between some keywords and the tourism volumes, and these keywords are eliminated. Based on the tourist psychology motivation angle and the Pearson correlation coefficient calculation results, as shown in **Table 1**, finally, choose the keyword with Pearson correlation coefficient above 0.3 as experimental data.

Using the standard scale we set as the limit, the abnormal data values in the search volume of the 5 keywords with the most relevant correlation are processed. When the search volume of keywords is lower or higher than the standard scale, many researchers will adopt the method of directly eliminating such abnormal data, but this will also cause excessive cleaning damage to the data. According to the seasonal characteristics of the tourism industry itself, this paper averages the other annual data of the month in which the data of the

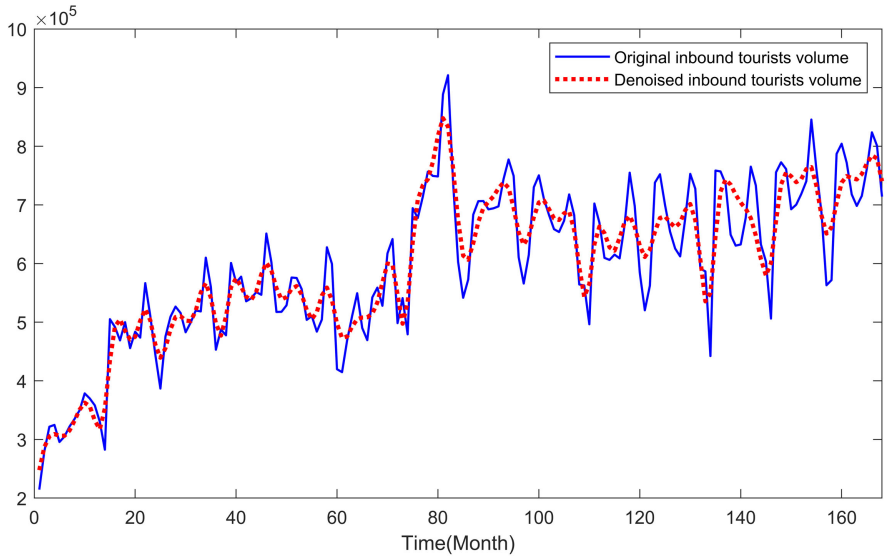


Figure 4. Tourism volumes sequence noise reduction.

Table 1. Search query correlation coefficient of Google keyword.

Keyword	Correlation Coefficient
Flight to Shanghai	0.597
Shanghai weather	0.128
Shanghai traffic	−0.420
Weather in Shanghai China	0.348
Shanghai tourist attractions	−0.183
China time Shanghai	0.656
Shanghai flight	0.479
Shanghai scene	−0.292
Shanghai hotel booking	0.024
Bund Shanghai	0.393
Shanghai cuisine	−0.312
Shanghai food	0.178
Shanghai airport	−0.272
Shanghai airport arrivals	0.045
Shanghai restaurant	0.124
Shanghai visa	0.053

abnormal data is taken to ensure the integrity of the data information.

The basic method of GDFM index composite is to use the weighted idea to sum up the common components of the variables. First, use the variance contribution rate to determine the number of factors, and then calculate the common components of the multidimensional stationary search data s_{it}, and finally add the search index.

$$s_{it} = \sum_{i=1}^{n} b_{it}\left(L^{k}\right) g_{nt} \tag{9}$$

n is the determined number of factors, L is the lag operator, and k is the number of lag operators. According to the Forni that when $n = 4$ and $k = 5$, the model works best [35].

After processing the abnormal data of the keyword search volume, the effect of the abnormal value on the authenticity of the search index is reduced. However the tourism is a relatively relevant industry, emergencies or other urban activities not related to tourism have a certain impact on tourism search, which has led to a significant increase or decrease in the overall network search of tourist destinations over time. For example, the SARS incidents in 2003 which affected the web search of tourist destinations. Abnormal data processing is for individual data, as shown in **Figure 5**. We still need to optimize the predictive power of the search index by simple noise reduction processing.

It can be seen that the composite search index does reflect the trend of tourist volumes and shows a certain lead time, as shown in **Figure 6**, which is also consistent with people's behavioral motives. At the peak of the tour, people seem more willing to use the search data in advance to help them make travel decisions.

4.3. Evaluating Indicator

In order to investigate the prediction ability of the established model from different angles, this paper selects the mean absolute percentage error (MAPE), the mean square error (MSE), the mean absolute error (MAE) and the fitting coefficient R^2 as the evaluation indicators of the model from different angles. Measure the predictive affect of the model. Among them, $\overline{x}_i$ represents the model simulation output value, that is, the predicted tourist volume; x_i represents the actual number of visitors, and n is the number of test data.

The fitting coefficient R^2 represents the degree of fitting of the predicted value curve to the actual value curve. The value of R^2 can measure the fitting ability of

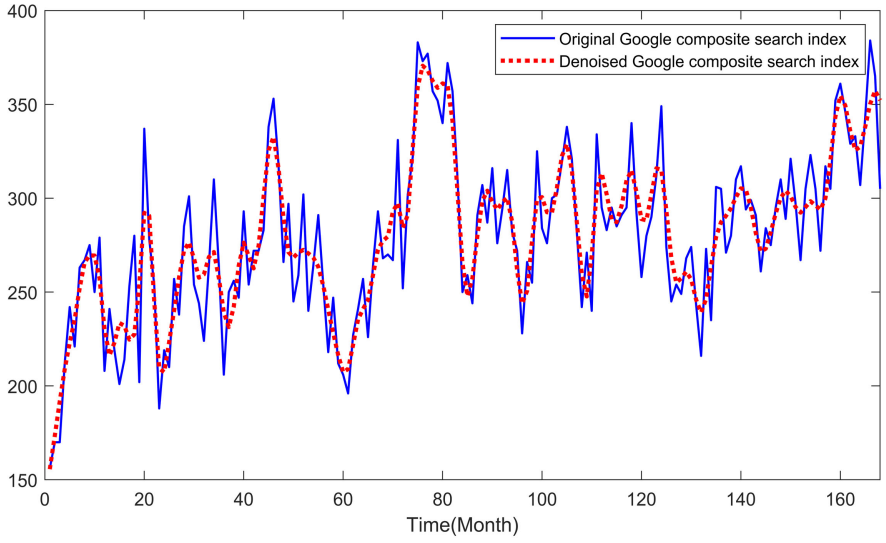


Figure 5. Search of key word index noise reduction.

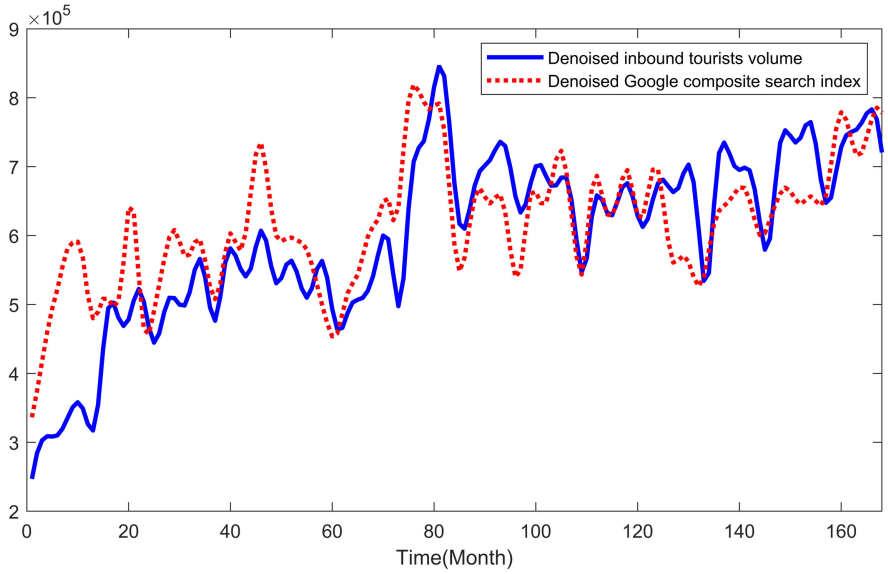


Figure 6. Fitting of tourism volumes sequence and composite search index.

the model, $R^2 \in [0,1]$, and the closer the value of R^2 is to 1, the model is explained the stronger the fitting ability.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2} \tag{10}$$

The MAE measures the accuracy of the prediction by calculating the difference between the predicted value and the true value data. The smaller value of MAE we have the higher prediction accuracy.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}_i| \tag{11}$$

MAPE is an evaluation criterion used to explain the relative error of the prediction model, which can well evaluate the prediction ability of the model.

$$\text{MAPE} = \frac{\sum_{i=1}^{n}|x_i - \bar{x}_i|/x_i}{n}100\% \tag{12}$$

RMSE represents the square root of the ratio of the sum of the predicted value to the true value and the ratio of the experimental number, used to estimate the degree of deviation between the predicted value and the true value.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_i)^2} \tag{13}$$

4.4. EEMD Noise Reduction

In the forecast research of tourism demand in recent years, there are many combined forecasting models, especially the combination of network search index and time series model. However, these studies lack the intensity of noise

processing. The modal decomposition of the tourist volumes time series and the keyword search index allows the Shanghai inbound tourist volumes and the keyword search index to obtain a uniformly distributed decomposition scale, so that smooth the interference of abnormal events.

The core idea of EEMD is to add Gaussian white noise to the signal and perform ensemble averaging. The two important parameters of EEMD are the ratio k of white noise to the standard deviation of the original signal amplitude, and the average number of times M, however there is no specific calculation method for the values of k and M. Combined with the experience of the researchers and the experiments in this paper, and for the data characteristics, we take k = 0.2 and M = 100 as the benchmark experimental values, and then adjust them continuously in the experiment to get the best EEMD model with the best decomposition effect.

When performing the EEMD test of the first iteration, the Gaussian white noise sequence $f_{(t)}$ was added to the Shanghai inbound tourist volumes time series $P_{(t)}$ and the keyword search index $I_{(t)}$, and n trials were performed. After that, the nth pending tourism volumes time series and the network search index sequence are obtained.

$$P_{n(t)} = p_{(t)} + lf_{n(t)} \tag{14}$$

$$I_{n(t)} = I_{(t)} + lf_{n(t)} \tag{15}$$

In further experiments, the EEMD parameter values continuously adjusted. The decomposed tourist volumes time series and the search index sequence contain 6 IMF components and one residual. The amplitude and fluctuation of the IMF component are different. It can also be seen from the figure that the amplitude and fluctuating frequency of the first IMF are always the largest and the wavelength is relatively short. The first IMF component obtained after decomposition is removed as noise, and the remaining IMF_2, IMF_3, IMF_4, IMF_5, IMF_6 and smooth trend residuals are summed as an experimental sequence of tourist volumes and keyword search respectively.

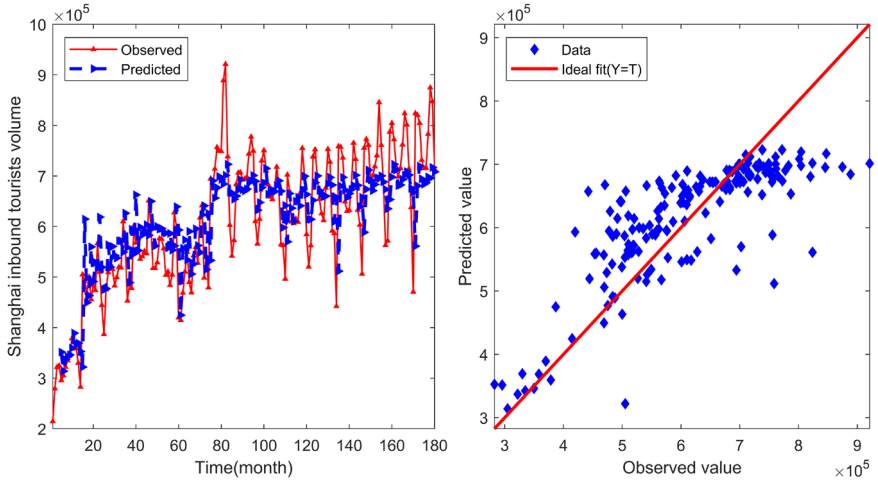
4.5. DBN Forecast

We are required to determine the number of hidden layer nodes and the number of hidden layers in the DBN. There is not any exact rule about the number of nodes in the DBN input layer and hidden layer. Basing on the experience of researchers, this paper sets the number of layers of DBN to N = 3 and the number of neurons is set at intervals of 5 for each hidden layer. The number of nodes in the layer is taken as an integer in [100, 1000], and then each additional layer of hidden layer is added to determine the optimal value of the number of neurons in the second hidden layer, and the experiment is repeated to achieve the highest accuracy. It component RBM of the DBN needs to optimize the feature extraction ability through training. Therefore, the DBN also needs a weight; the weight can determine the influence factor of the maximum probability of the training

sample. We define the learning rate of the DBN model to 0.1, the number of iterations. Set to200, by repeating the training, it is determined that the final DBN structure is a 3-layer RBM composition, and the number of hidden neurons in the first layer and the second layer is 20, 15, respectively.

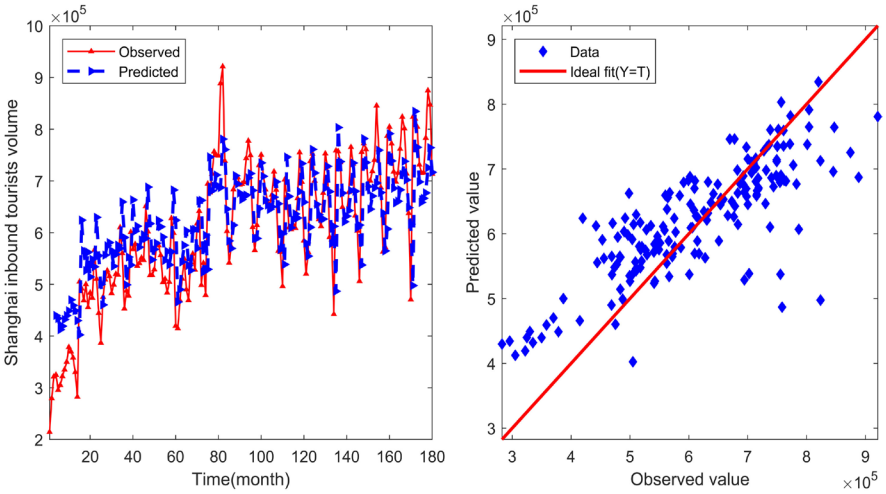
In this paper, the established DBN model is used for prediction and compared with support vector machine, ARIMA, GM(1,1), fuzzy time series and cubic exponential smoothing model. The group (a) of Figures 6-11 shows the fit between the predicted value and the actual value, and the group (b) of Figures 7-12 shows the degree of dispersion between the predicted value and the observed value.

It can be seen from the comparison of the (a) group images of the six different



(a) (b)

Figure 7. Forecasts of the monthly Shanghai inbound tourist volume using GDFM-ARIMA: (a) the predicted series; and (b) the scatter of the predicted series.



(a) (b)

Figure 8. Forecasts of the monthly Shanghai inbound tourist volume using GDFM-GM(1,1): (a) the predicted series; and (b) the scatter of the predicted series.

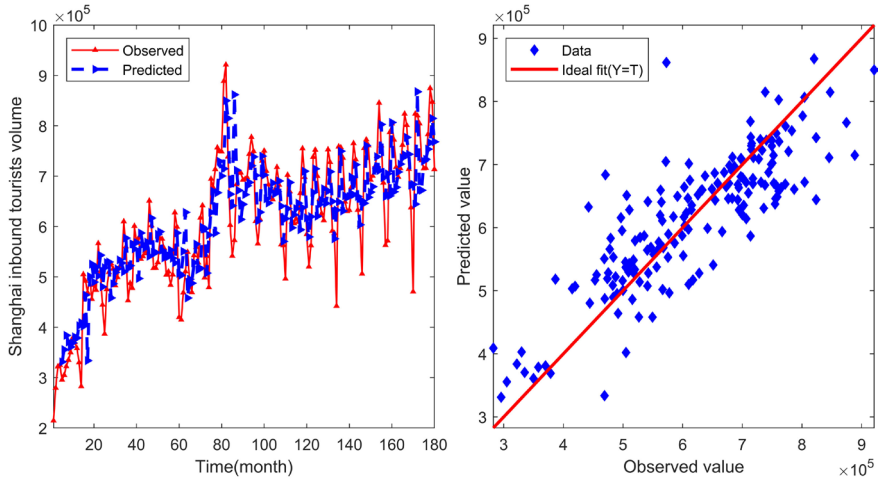


Figure 9. Forecasts of the monthly Shanghai inbounds tourist volume using GDFM-FTS: (a) the predicted series; and (b) the scatter of the predicted series.

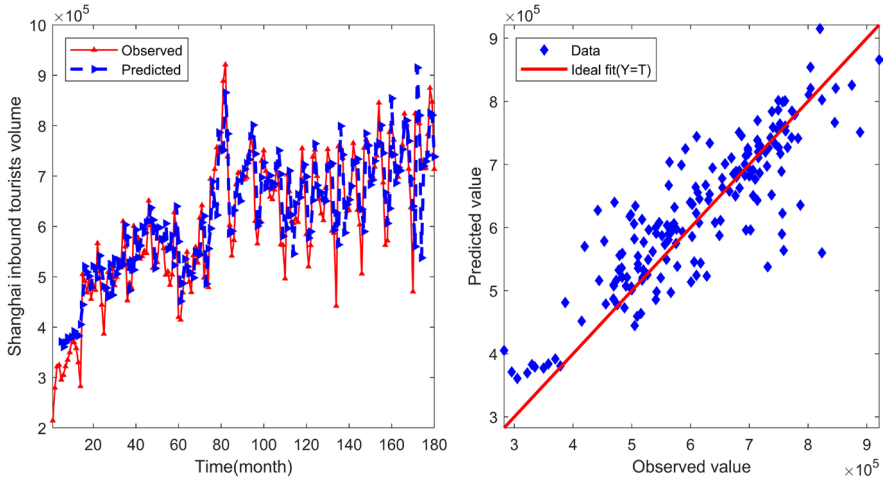


Figure 10. Forecasts of the monthly Shanghai inbound tourist volume using GDFM-SVM: (a) the predicted series; and (b) the scatter of the predicted series.

prediction models that the predicted values of the ARIMA model show obvious convergence characteristics, and ARIMA has greater limitations in dealing with non-stationary time series. The prediction ability of the GM(1,1) model usually shows a large volatility, which is affected by the smoothness of the tourist volumes data series, the prediction effect of the GM(1,1) model seems to be less than ideal. FTS usually optimizes the uncertainty of data and solves fuzzy problems. When forecasting separately it often does not really work out well. SVM is a widely used forecasting model in tourism demand forecasting in recent years. However, the kernel function selection of SVM model is a very difficult problem, and it is computationally complex and often sensitive to data loss. The cubic exponential smoothing is based on an exponential smoothing model and a quadratic exponential smoothing model. It is commonly used in China's domestic

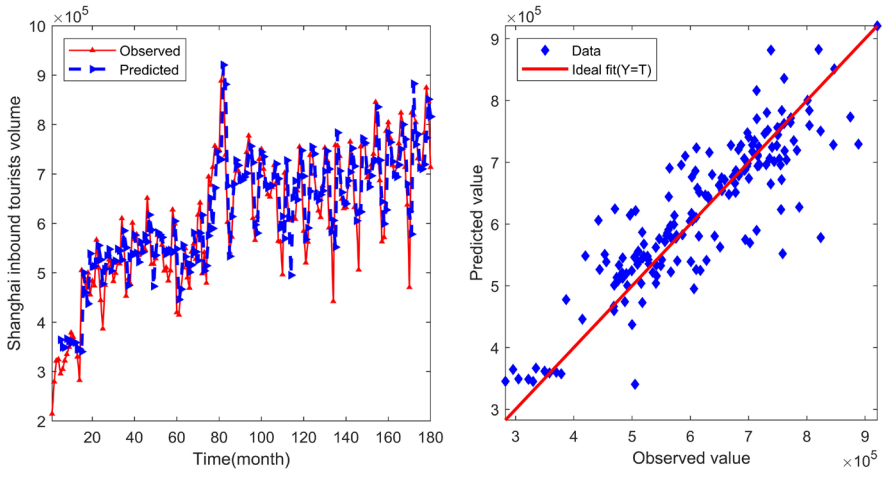


Figure 11. Forecasts of the monthly Shanghai inbound tourist volume using GDFM-CES: (a) the predicted series; and (b) the scatter of the predicted series.

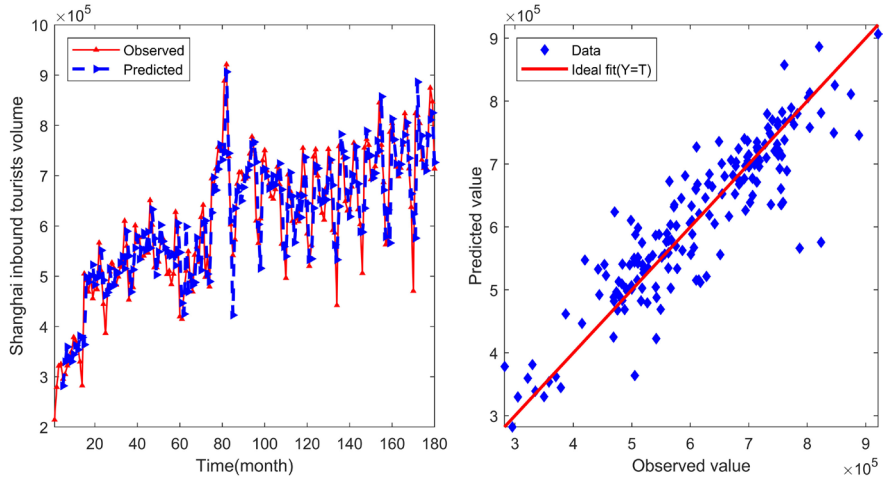


Figure 12. Forecasts of the monthly Shanghai inbound tourist volume using GDFM-DBN: (a) the predicted series; and (b) the scatter of the predicted series.

tourism demand forecast because it has good predictive ability for seasonal time series. From the figure, the fitting effect of the cubic exponential smoothing and DBN is relatively trustworthy.

The group image of (b) is a good representation of the degree of dispersion between the ideal prediction and the actual predictions of the six different models. The discrete trend of the DBN model is the most stable and the least discrete, that is, the predicted value of the DBN model is closely surrounding the actual value. Combining the final results of the (a) group image with the (b) group image, we can conclude that the DBN has better predictive power than other models.

4.6. Comparison of Different Forecasting Methods

The comparison experiment was set up in two groups. The first group predicted

the tourist volumes without adding the composite search index, and the second group joined the Google search index to predict the tourist volumes. The results of both groups were evaluated by MAE, MAPE, RMSE, and R^2 as shown in Table 2. The MAE, MAPE, RMSE, and R^2 values predicted by the first group of ARIMA models were 10.16828, 0.14032, 1.27639, and 0.62679, respectively; the MAE, MAPE, RMSE, and R^2 values predicted by GM(1,1) were 8.47406, 0.1159, 1.21326, and 0.63701, respectively; The MAE, MAPE, RMSE, and R^2 values predicted by fuzzy time series are 7.77446, 0.11376, 0.99984, and 0.70957, respectively; the MAE, MAPE, RMSE, and R^2 values predicted by SVM are 7.95296, 0.11417, 1.13246, and 0.74384, respectively; cubic exponential smoothing model prediction The MAE, MAPE, RMSE, and R^2 values were 7.31596, 0.10397, 0.99589, and 0.76065, respectively; the MAE, MAPE, RMSE, and R^2 values predicted by DBN were 5.8507, 0.08417, 0.90995, and 0.80651, respectively. Obviously, the evaluation indicators of the DBN model are better than other models, especially the values of MAE and MAPE are nearly half smaller than ARIMA, and the R^2 value is also the closest to 1.

Looking at the results of the second set of experiments, after joining the Google keyword search index, we used six different models to predict the tourist volumes as shown in Table 3. The MAE, MAPE, RMSE, and R^2 values predicted by ARIMA after adding the search index are 10.40777, 0.15558, 1.30863, and 0.64662, respectively; the MAE, MAPE, RMSE, and R^2 values predicted by GM(1,1) are 7.03853, 0.09941, 0.85096, and 0.64336 respectively; MAE, MAPE, RMSE, and R^2 values predicted by fuzzy time series are 8.02626, 0.10963, 1.05467, and 0.72485, respectively; the MAE, MAPE, RMSE, and R^2 values predicted by SVM are 6.28613, 0.09048, 1.00487, and 0.779, respectively; The MAE, MAPE, RMSE, and R^2 values predicted by the smoothing model were 5.99157, 0.084653, 0.91868, and 0.78252, respectively; the MAE, MAPE, RMSE, and R^2 values predicted by DBN were 5.04885, 0.071167, 0.65961, and 0.82518, respectively. Compared with the results of the first set of experiments, it can be found that after adding the Google keyword search index, the MAE, MAPE, and RMSE values of GM(1,1), SVM, CES, and DBN are smaller than those when the search index is not added, also improved in R^2. Obviously, the DBN model still maintains the most predictive performance, and the indicators are ahead of other forecasting models. This shows that the DBN model does show a good forecasting

Table 2. Forecasting performance evaluation of six models.

	MAE (10^4)	MAPR	RMSE (10^5)	R^2
ARIMA	10.16828	0.14032	1.27639	0.62679
GM(1,1)	8.47406	0.1159	1.21326	0.63701
FTS	7.77496	0.11376	0.99984	0.70957
SVM	7.95296	0.11417	1.13246	0.74384
CES	7.31596	0.10397	0.99589	0.76065
DBN	5.8507	0.08417	0.90995	0.80651

Table 3. Forecasting performance evaluation of six models with search index.

	MAE (10^4)	MAPR	RMSE (10^5)	R^2
Index-ARIMA	10.40777	0.15558	1.30863	0.64662
Index-GM(1,1)	7.03853	0.09941	0.85096	0.64336
Index-FTS	8.02626	0.10963	1.05467	0.72485
Index-SVM	6.28613	0.09048	1.00487	0.779
Index-CES	5.99157	0.084653	0.91868	0.78252
Index-DBN	5.04885	0.071167	0.65961	0.82518

effect in terms of tourist volumes forecasting, and the Google search index can indeed optimize the forecast of tourist volumes.

4.7. Stability Test and Granger Causality Test

The granger causality test is to prove the effectiveness of Google keyword search data on tourism demand forecasting model. Before conducting the Granger causality test, we must first ensure that the tourist volumes sequence and the Google search index sequence are stationary. According to the unit root test results, the Google search index is stationary, at the 1% significance level (**Table 4**). The tourist volumes sequence shows the non-stationary, state during the same period, so the differential processing has to be done. Under the second-order differential level, the tourist volumes sequence is stationary (**Table 5**). Further, Granger causality test between variables is shown in **Table 6**. According to the results of the Granger causality test, Google search data is the cause of tourists' travel behavior. Then it proves that the Google search index has predictive ability for inbound tourism in Shanghai. Therefore, it is feasible to use the search data to predict the tourism volumes. The tourism volumes sequence is recorded as Y_t, and the keyword search index is recorded as I_t.

Table 4. Augmented Dickey-Fuller unit root test on Google search index.

		t-statistic	Prob
Augmented Dickey-Fuller test statistic		−4.447879	0.0002
Test critical values:	1% level	−3.467205	
	5% level	−2.877636	
	10% level	−2.575430	

Table 5. Augmented Dickey-Fuller unit root test on tourism volumes (2nd difference).

		t-statistic	Prob
Augmented Dickey-Fuller test statistic		−12.42683	<0.0001
Test critical values:	1% level	−3.469933	
	5% level	−2.878829	
	10% level	−2.576067	

Table 6. Standard granger causality tests.

null hypothesis	F-statistic	Prob
Y_t does not Granger cause I_t	7.39027	0.0001
I_t does not Granger cause Y_t	3.06659	0.0295

5. Conclusions

Network technology is constantly upgrading, and has achieved good popularity, becoming an indispensable part of people's daily lives. With the advent of the 5G era, web search may penetrate deeper into our daily lives, especially in terms of travel. The EEMD decomposition method adopted in this paper overcomes the large noise defects in the traditional index composite, making the keyword search index play the most important role in tourism demand forecasting. However, we have to admit that the selection of keywords is something we need to explore further, although it is now possible to use high-speed computers to extract keywords, and such methods will greatly improve the accuracy of keywords, but this technology extremely high hardware requirements and therefore no universality. Accurately finding keywords that best reflect the motivation of tourist guests will further optimize the tourism demand forecasting model. Five comparative forecasting models selected in this paper are widely used in tourism demand forecasting. The maturity of artificial intelligence technology will also bring new opportunities and challenges to tourism demand forecasting.

Tourism is a comprehensive industry that will not only be affected by force major such as weather, natural disasters, but also subjective factors such as politics, economy, culture and even religion. In the traditional forecast of tourism demand, the research using quantitative methods accounts for the majority, which also neglects the tourism behavior caused by people's subjective consciousness to some extent, and the keyword search data, reflects the subjective behavior of people. Although the keyword information has been applied to some extent, it is not deep enough. Therefore, accurate extraction of keywords and qualitative analysis is another challenging problem that we need to work hard to solve.

Acknowledgements

This research is supported by Self determined Research Funds of CCNU from the Colleges' Basic Research and Operation of MOE under Grant No. CCNU19ZN024.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Clerides, S. and Adamou, A. (2010) Prospects and Limits of Tourism-Led Growth:

The International Evidence. *Review of Economic Analysis*, **2**, 287-303. https://doi.org/10.2139/ssrn.1495747

[2] Fesenmaier, D., Xiang, Z., Pan, B. and Law, R. (2011) A Framework of Search Engine Use for Travel Planning. *Journal of Travel Research*, **50**, 587-601. https://doi.org/10.1177/0047287510385466

[3] Choi, H. and Varian, H. (2009) Predicting the Present with Google Trends. Technical Report, Google Inc. https://doi.org/10.2139/ssrn.1659302

[4] Chan, C., Witt, S., Lee, Y. and Song, H. (2010) Tourism Forecast Combination Using the CUSUM Technique. *Tourism Management*, **31**, 891-897. https://doi.org/10.1016/j.tourman.2009.10.004

[5] Song, H. and Li, G. (2008) Tourism Demand Modeling and Forecasting: A Review of Recent Research. *Tourism Management*, **29**, 203-220. https://doi.org/10.1016/j.tourman.2007.07.016

[6] Shen, S., Li, G. and Song, H. (2011) Combination Forecast of International Tourism Demand. *Annals of Tourism Research*, **38**, 72-89. https://doi.org/10.1016/j.annals.2010.05.003

[7] Peng, B., Song, H. and Crouch, G. (2014) A Meta-Analysis of International Tourism Demand Forecasting and Implications for Practice. *Tourism Management*, **45**, 181-193. https://doi.org/10.1016/j.tourman.2014.04.005

[8] Xiao, Y., Liu, J., Xiao, J., Hu, Y., Bu, H. and Wang, S.Y. (2015) Application of Multiscale Analysis-Based Intelligent Ensemble Modeling on Airport Traffic Forecast. *Transportation Letters*: *The International Journal of Transportation Research*, **7**, 73-79. https://doi.org/10.1179/1942787514Y.0000000035

[9] Xiao, Y., Liu, Y., Liu, J.J., Xiao, J. and Hu, Y. (2016) Oscillations Extracting for the Management of Passenger Flows in the Airport of Hong Kong. *Transportmetrica A*: *Transport Science*, **12**, 65-79. https://doi.org/10.1080/23249935.2015.1099576

[10] Xiao, Y., Liu, J.J., Hu, Y., Wang, Y.F., Lai, K.K. and Wang, S.Y. (2014) A Neuro-Fuzzy Combination Model Based on Singular Spectrum Analysis for Air Transport Demand Forecasting. *Journal of Air Transport Management*, **39**, 1-11. https://doi.org/10.1016/j.jairtraman.2014.03.004

[11] Lim, C. and McAleer, M. (2001) Monthly Seasonal Variations: Asian Tourism to Australia. *Annals of Tourism Research*, **28**, 68-82. https://doi.org/10.1016/S0160-7383(00)00002-5

[12] Lim, C. and McAleer, M. (2002) Time Series Forecasts of International Travel Demand for Australia. *Tourism Management*, **23**, 389-396. https://doi.org/10.1016/S0261-5177(01)00098-X

[13] Martin, A. and Witt, F. (1989) Forecasting Tourism Demand: A Comparison of the Accuracy of Several Quantitative Methods. *International Journal of Forecasting*, **5**, 1-13. https://doi.org/10.1016/0169-2070(89)90059-9

[14] Gounopoulos, D., Petmezas, D. and Santamaria, D. (2012) Forecasting Tourist Arrivals in Greece and the Impact of Macroeconomic Shocks from the Countries of Tourists' Origin. *Annals of Tourism Research*, **39**, 641-666. https://doi.org/10.1016/j.annals.2011.09.001

[15] Chu, F.L. (2009) Forecasting Tourism Demand with ARMA-Based Methods. *Tourism Management*, **30**, 740-751. https://doi.org/10.1016/j.tourman.2008.10.016

[16] Chen, J., Chen, Z.X., Xing, L. and Fu, X.D. (2005) Forecasting of Yunnan's International Tourism Demand Based on BP Neural Network. *Journal of Kunming Teachers College*, **27**, 89-91.

[17] Goh, C. and Law, R. (2003) Incorporating the Rough Sets Theory into Travel Demand Analysis. *Tourism Management*, **24**, 511-517. https://doi.org/10.1016/S0261-5177(03)00009-8

[18] Samagaio, A. and Wolters, M. (2010) Comparative Analysis of Government Forecasts for the Lisbon Airport. *Journal of Air Transport Management*, **16**, 213-217. https://doi.org/10.1016/j.jairtraman.2009.09.002

[19] Nguyen, L.T., Shu, H.M., Huang, F.Y. and Hsu, M.B. (2013) Accurate Forecasting Models in Predicting the Inbound Tourism Demand in Vietnam. *Journal of Statistics and Management Systems*, **16**, 25-43. https://doi.org/10.1080/09720510.2013.777570

[20] Chen, K.Y. and Wang, C.H. (2007) Support Vector Regression with Genetic Algorithms in Forecasting Tourism Demand. *Tourism Management*, **28**, 215-226. https://doi.org/10.1016/j.tourman.2005.12.018

[21] Burger, C., Dohnal, M., Kathrada, M. and Law, R. (2001) A Practitioners Guide to Time Series Method for Tourism Demand Forecasting: A Case Study of Durban, South Africa. *Tourism Management*, **22**, 403-409. https://doi.org/10.1016/S0261-5177(00)00068-6

[22] Alvarez-Diaz, M., Mateu-Sbert, J. and Rosselló-Nadal, J. (2009) Forecasting Tourist Arrivals to Balearic Island Using Genetic Programming. *International Journal of Computational Economics and Econometrics*, **1**, 65-75. https://doi.org/10.1504/IJCEE.2009.029153

[23] Jackman, M. and Naitram, S. (2015) Nowcasting Tourist Arrivals in Barbados: Just Google It! *Tourism Economics*, **21**, 1309-1313. https://doi.org/10.5367/te.2014.0402

[24] Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S. and Brilliant, L. (2009) Detecting Influenza Epidemics Using Search Engine Query Data. *Nature*, **457**, 1012-1014. https://doi.org/10.1038/nature07634

[25] Askitas, N. and Zimmermann, K.F. (2009) Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly*, **55**, 107-120. https://doi.org/10.3790/aeq.55.2.107

[26] Beracha, E. and Wintoki, M.B. (2013) Forecasting Residential Real Estate Price Changes from Online Search Activity. *Journal of Real Estate Research*, **35**, 283-312.

[27] Da, Z., Engelberg, J. and Gao, P. (2011) In Search of Attention. *Journal of Finance*, **66**, 1461-1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

[28] Bangwayo-Skeete, P.F. and Skeete, R.W. (2015) Can Google Data Improve the Forecasting Performance of Tourist Arrivals? Mixed-Data Sampling Approach. *Tourism Management*, **46**, 454-464. https://doi.org/10.1016/j.tourman.2014.07.014

[29] Yang, X., Pan, B., Evans, J.A. and Lv, B. (2015) Forecasting Chinese Tourist Volume with Search Engine Data. *Tourism Management*, **46**, 386-397. https://doi.org/10.1016/j.tourman.2014.07.019

[30] Theologos, D., Eleni, M. and Bing, P. (2018) Google Trends and Tourists' Arrivals: Emerging Biases and Proposed Corrections. *Tourism Management*, **66**, 108-120. https://doi.org/10.1016/j.tourman.2017.10.014

[31] Yassin, I.M., Zabidi, A., Salleh, M.K.M. and Khalid, N.E.A. (2013) Malaysian Tourism Interest Forecasting Using Nonlinear Auto-Regressive (NAR) Model. *Proceedings of IEEE* 3*rd International Conference on System Engineering and Technology*, Shah Alam, 19-20 August 2013, 32-36. https://doi.org/10.1109/ICSEngT.2013.6650138

[32] Kennedy, A.F. and Hauksson, K.M. (2012) Global Search Engine Marketing:

Fine-Tuning Your International Search Engine Results. Que Publishing, Indianapolis.

[33] Huang, N.E., Shen, Z., Long, S.R., *et al*. (1998) The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings of the Royal Society A*: *Mathematical*, *Physical and Engineering Sciences*, **454**, 903-995. https://doi.org/10.1098/rspa.1998.0193

[34] Chen, C.F., Lai, M.C. and Yeh, C.C. (2012) Forecasting Tourism Demand Based on Empirical Mode Decomposition and Neural Network. *Knowledge-Based Systems*, **26**, 281-287. https://doi.org/10.1016/j.knosys.2011.09.002

[35] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000) The Generalized Dynamic-Factor Model: Identification and Estimation. *Review of Economics and Statistics*, **82**, 540-554. https://doi.org/10.1162/003465300559037

[36] Amstad, M. and Potter, S. (2009) Real Time Underlying Inflation Gauges for Monetary Policymakers. FRB of New York Staff Report, 420. https://doi.org/10.2139/ssrn.1532280

[37] Forni, M., Giannone, D., Lippi, F. and Reichlin, L. (2004) Opening the Black Box Structural Factor Models versus Structural VARS. Discussion Papers of CEPR.

[38] Hinton, G.E., Osindero, S. and The, Y.W. (2006) A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, **18**, 1527-1554. https://doi.org/10.1162/neco.2006.18.7.1527

[39] Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. and Kingsbury, B. (2012) Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, **29**, 82-97. https://doi.org/10.1109/MSP.2012.2205597

[40] Hinton, G. (2002) Training Products of Experts by Minimizing Contrastive Divergence. *Neural Computation*, **14**, 1771-1800. https://doi.org/10.1162/089976602760128018

Scientific Research Publishing

Engineering, 2020, 12, 216-227
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

The Mathematical Model of Seed Movement on a Concave Profile Rib

Khayrullo Sharipov, Khamit Akhmedxodjayev, Mukhammadjon Tojiboyev, Olimjon Sarimsakov*

Department of Technology of Primary Processing of Natural Fibers, Namangan Institute of Engineering and Technology, Namangan, Uzbekistan
Email: *Olimjon5008@mail.ru

How to cite this paper: Sharipov, K., Akhmedxodjayev, K., Tojiboyev, M. and Sarimsakov, O. (2020) The Mathematical Model of Seed Movement on a Concave Profile Rib. *Engineering*, **12**, 216-227.
https://doi.org/10.4236/eng.2020.123017

Received: December 3, 2019
Accepted: March 27, 2020
Published: March 30, 2020

Abstract

The differential equations of movement along the concave profile of the grate, consisting of three broken lines, are integrated on Maple 9.5 under initial conditions, using separate functions, and graphs of the dependence of movement and speed over time are presented. The graphs show the patterns of change in displacement and speed at different angles, friction coefficient of seeds along grate with a broken line of a concave profile.

Keywords

Ginning of Cotton, Saw Gin Stand, Roll Box, Seed Roll, Rib, Concave, Cotton Seed

1. Introduction

Currently, the quality of fiber obtained at cotton processing enterprises depends on the efficient operation of machines operating in the direct technological process. Each process is to some extent important in obtaining high-quality fiber. The main technological process in the production of fiber at the enterprise is the process of ginning (separation of fiber from seed). The cotton, peeled from small and large litter in the cleaning workshops, is fed to the main machine of the genie shop with saw gin. The cotton enters the working chamber of the gin and is hooked with the teeth of the saws rotating next to the seed comb, leading to the grate. In the working chamber, the cotton flies hooked with the teeth saw, hook the other cotton flies and create a raw roller. The seed roll rotates in the opposite direction of the saw cylinder, and provides a continuous supply of cotton fiber to the teeth of the saws. Gin problems were previously studied by many researchers [1]-[7].

There are also known attempts by the authors of mathematical modeling of

various stages of the technology of primary processing of cotton, which are published both in local and foreign scientific journals [8] [9] [10] [11] [12].

The authors of the article conducted a series of studies to improve the working elements of gin. The aim of the study is to create the possibility of timely exit of bare seeds from the working chamber of the saw gin by creating grooves in the grates, creating a device that performs this process, determining the technological dimensions that ensure efficient operation, as well as introducing the device into production.

The selection of the optimal structural and technological parameters of the new grate is a crucial stage of research. The use of mathematical methods in research planning, in contrast to traditional methods of research calculations, makes it possible to determine the influence of each factor together influencing several parameters on optimization parameters. As a result of this, a mathematical model of the studied object will be obtained with a relatively small multiplicity of tests. This model is also used for decision making.

Investigation of the movement of single and systemic seeds along grates with a concave profile.

1) The movement of a single seed.

To describe the movement of the seed, it is necessary to determine the concavity of the grate groove. Typically, this concavity profile consists of a curve, the shape of which should provide the minimum resistance force and the process of maximum separation of the fiber from the seed when the seeds move. Suppose the equation of the concavity profile parameter, and the shape is expressed by a curved line in the plane (Figure 1).

In this case, its curvilinear motion will be expressed by the following equations:

$$\begin{cases} m \cdot \dfrac{\mathrm{d}^2 S}{\mathrm{d}t^2} = \sum F_{n\tau} \\ \dfrac{m\upsilon^2}{\rho} = \sum F_{nu}^{(a)} + N \end{cases} \tag{1}$$

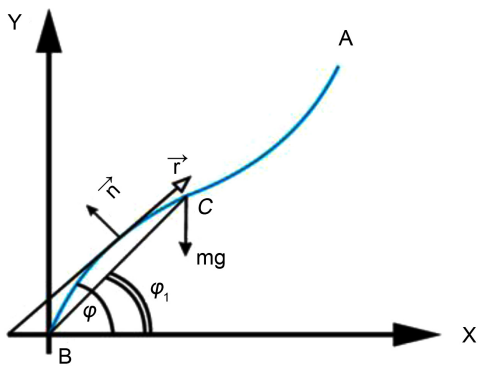


Figure 1. The movement of the seed along the rib with a concave profile.

There is $S = S(t)$—the law of movement of the seed in concavity;

ρ—radius of curvature;

$F_{n\tau}$ and F_{nu}—projections of external forces to the curve in the tangent and normal directions;

N—normal reaction force affecting the mass.

Consider the equation in the polar coordinate system. We consider the equation of the curve of the line in the coordinate of the pole systems. The law of motion of the seed is written as follows:

$$\left.\begin{aligned} m\frac{\mathrm{d}^2 S}{\mathrm{d}t_i^2} &= T \\ N &= \frac{mv^2}{p} - mg\cos\psi \end{aligned}\right\} \tag{2}$$

ψ—the angle formed by the tangent to the curve of the line with the axis *ox*, this angle is expressed as follows:

$$\left.\begin{aligned} \mathrm{tg}\,\psi &= \frac{\frac{\mathrm{d}y}{\mathrm{d}\varphi}}{\frac{\mathrm{d}x}{\mathrm{d}\varphi}} = \frac{\dot{r}\sin\varphi + r\cos\varphi}{\dot{r}\cos\varphi - r\sin\varphi} \\ \sin\psi &= \frac{\mathrm{tg}\,\psi}{\sqrt{1+\mathrm{tg}^2\psi}} = \frac{\dot{r}\sin\varphi + \cos\varphi}{\sqrt{r^2+\dot{r}^2}} \\ \cos\psi &= \frac{\dot{r}\cos\varphi - r\sin\varphi}{\sqrt{r^2+\dot{r}^2}} \end{aligned}\right\} \tag{3}$$

For the radius of curvature we obtain the following expression

$$\rho = \frac{\left(r^2 + r^{12}\right)^{312}}{r^2 + 2r^{12} - rr^1} \tag{4}$$

In Equation (2), the letter indicates the sum of the tangential forces affecting the seed. In this direction, the projection of gravity and the Coulomb friction force in the direction, *i.e.*

$$T = fN + mg\sin\psi$$

If we write Equation (2) taking into account the expressions and for, we obtain the following equation:

$$\begin{aligned} m\ddot{S} = {} & \frac{mg}{\sqrt{r^2+r^{12}}}\left[\dot{r}\cos\varphi + r\cos\varphi + f\left(\dot{r}\cos\varphi - r\sin\varphi\right)\right] \\ & - fm\dot{S}^2\frac{r^2 + 2r^{12} - r\ddot{r}}{\sqrt{r^2+r^{12}}} \end{aligned} \tag{5}$$

In this equation using equality $\dot{S} = \dot{\varphi}\sqrt{r^2+r^{12}}$, for the angle $\varphi = \varphi(t)$ we can write the following nonlinear differential equation of the second order:

$$\begin{aligned} & \left[-\ddot{\varphi}\sqrt{r^2+r^{12}} + \dot{\varphi}^2\frac{r^1\left(r+r^{11}\right)}{\sqrt{r^2+r^{12}}}\right] \\ & = \frac{g\left(\dot{r}\sin\varphi + r\cos\varphi\right) + f\left(\dot{r}\cos\varphi - r\sin\varphi\right)}{\sqrt{r^2+\dot{r}^2}} - f\dot{\varphi}^2\left(r^2 + 2\dot{r}^2 - \dot{r}^1\sqrt{r^2+\dot{r}^2}\right) \end{aligned}$$

The initial equations of this equation $\varphi(0)=\varphi_0$ and $\dot{\varphi}(0)=\dot{\varphi}_0$ (6) the differential equation are integrated numerically under the above initial conditions.

The particular case (Figure 2):

The lengths of the convex grate l_1, l_2, l_3 are replaced by three straight lines *BD*, *DC* and *CA*.

Enter the direct Cartesian coordinates for the moving seed in each plot

$$\left.\begin{array}{l} x_1=(l_1-S_1)\cos a_1 \\ y_1=(l_1-S_1)\sin a_1 \end{array}\right\} \quad 0<S_1<l_1$$
$$\left.\begin{array}{l} x_2=l_1\cos a_1+S_2\cos a_2 \\ y_2=l_2\sin a_1+S_2\sin a_1 \end{array}\right\} \quad 0<S_2<l_2$$
$$\left.\begin{array}{l} x_3=l_1\cos a_1+l_2\cos a_2+S_3\cos a_3 \\ y_3=l_1\sin a_1+l_2\sin a_1+S_3\sin a_3 \end{array}\right\} \quad 0<S_3<l_3 \qquad (7)$$

We can write the differential equation of motion $S=S(t)$ of the path covered by the seed for each section as follows:

$$\begin{cases} m\ddot{S}_1=mg(\sin a_1-f\cos a_1) & 0<t<t_1 \qquad (8) \\ m\ddot{S}_2=mg(\sin a_2-f\cos a_2) & t_1<t<t_2 \qquad (9) \\ m\ddot{S}_3=mg(\sin a_3-f\cos a_3) & t_2<t<t_3 \qquad (10) \end{cases}$$

Equations (8)-(10) are integrated under the following initial conditions

$$\left.\begin{array}{ll} S_1(0)=0 & \dot{S}_1(0)=\vartheta_1 \\ S_2(t_1)=0 & \dot{S}_2(t_1)=\vartheta_2 \\ S_2(t_2)=0 & \dot{S}_2(t_2)=\vartheta_3 \end{array}\right\} \qquad (11)$$

Here t_1, t_2, t_3, v_2, v_3, ··· To do this, consider each period of time.

a) $0<t<t_1$—at this time, the seed moves only along the line *AB* and from Equation (2) under the conditions $S_1(0)=0$, $\dot{S}_1(0)=v_h$ the integral will be in the following form:

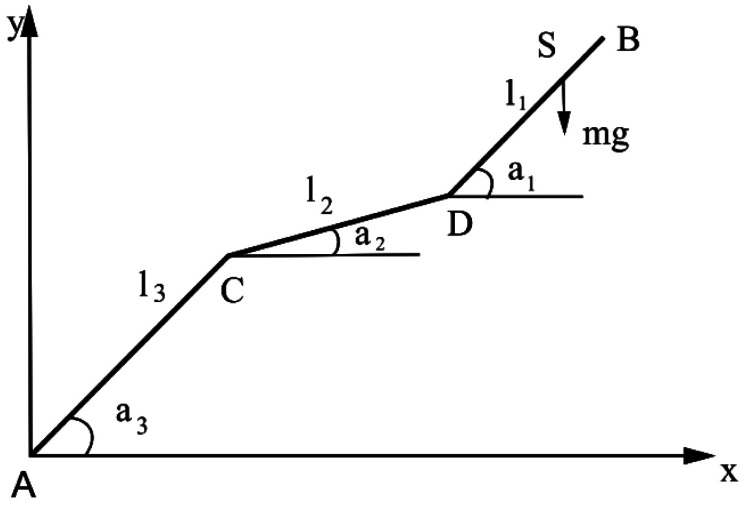


Figure 2. Profile view of a rib with 3 broken straight lines.

$$S_1 = C_1 \frac{t^2}{2} + v_h t \tag{12}$$

$$\left(C_1 = \sin a_1 - f \cos a_1\right)$$

Here $v_h = \sqrt{2gh}$, h—is the height from the center of gravity of the seed to the bulge of the grate.

In Equation (12), taking into account the equation $S_1 = l_1$, we obtain the equation with respect to t_1

$$C_1 t_1 + 2v_n t_1 - l_1 = 0$$

$$l_1 = c \cdot \frac{t^2}{2} + \upsilon_h \cdot t \Rightarrow \tag{13}$$

Here we define $t = t_1$.

As a result, we obtain the expression of speed v_1

$$v_1 = c_1 t_1 + v_h \tag{14}$$

Similarly, we obtain the following solutions for the second and third sections

$$\left.\begin{array}{l} S_2 = c_2 \dfrac{\left(t - t_1\right)^2}{2} + \vartheta_2 \left(t - t_1\right) \\ \vartheta_2 = \vartheta_1 \cos a_1 \\ c_2 = \sin a_2 - f \cos a_2 \end{array}\right\} \tag{15}$$

$$\left.\begin{array}{l} S_3 = c_3 \dfrac{t - t_2}{2} + \vartheta_3 \left(t - t_2\right) \\ \vartheta_3 = \vartheta_2 \cos a_2 \\ c_3 = \sin a_3 - f \cos a_3 \end{array}\right\} \tag{16}$$

2) Assume that the sections of the transition *AB*, *BC* and *CA* the convex grate are expressed by circular arcs (Figure 3).

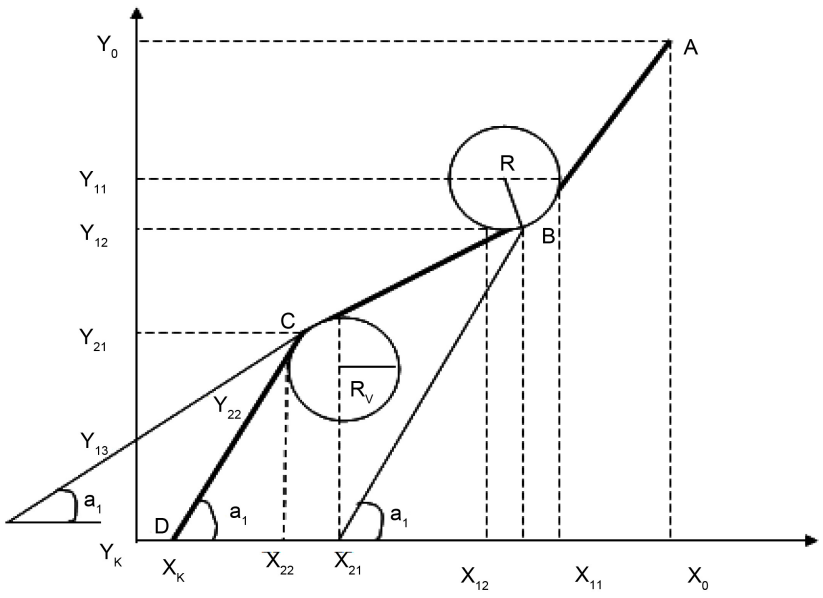


Figure 3. Coordinates of the characteristic points of the rib.

We write the equations of arcs in the following form

$$\left.\begin{aligned}(x-x_1)^2+(y-y_1)^2=R_1^2\\(x-x_2)^2+(y-y_2)^2=R_i\end{aligned}\right\} \tag{17}$$

Here x_i; y_i—coordinates of the centers of arcs;

R_i—radius of arcs.

Equations of straight lines, *AB*, *BC* and *CD*:

$$\left.\begin{aligned}y_1=k_1(x-x_{11})+y_{11}\\y_2=x_2(x-x_{12})+y_{12}\\y_3=x_3(x-x_{13})+y_{13}\end{aligned}\right\} \tag{18}$$

Points x_{11}, y_{11} lie on an arc, therefore

$$(x_{11}-x_1)^2+(y_{11}-y_1)^2=R_i^2 \tag{19}$$

In addition, the slope of the line *AB*

$$K_1=\frac{x_{11}-x_1}{\sqrt{(x_{11}-x_1)^2+(y_{11}-y_1)^2}}=\mathrm{tg}a_1 \tag{20}$$

If the angle coefficient K_1 is given, the coordinates x_{11} and y_{11} are also determined in Equations (18) and (20).

The coordinates x_{12} and y_{12}

$$(x_{12}-x_1)^2+(y_{12}-y_1)^2=R_i \tag{21}$$

$$K_2=\frac{x_{12}-x_1}{\sqrt{(x_{12}-x_1)^2+(y_{12}-y_1)}}=\mathrm{tg}a_2 \tag{22}$$

$$(x_{21}-x_2)^2+(y_{21}-y_2)^2=R_i^2 \tag{23}$$

$$K_3=\frac{x_{22}-x_2}{\sqrt{(x_{22}-x_2)^2+(y_{22}-y_2)^2}}=\mathrm{tg}a_3 \tag{24}$$

From each Equations (20)-(24) unknown coordinates are determined x_{11}, y_{11}; x_{12}, y_{12}; x_{21}, y_{21}; x_{22}, y_{22}. The equation for the convexity of the grate will be in the form

$$\begin{aligned}&y=K_1(x-x_{11})+y_{11} && x_0<x<x_{11}\\&y=y_1-\sqrt{R_i^2-(x-x_1)^2} && x_{11}<x<x_{12}\\&y=K_2(x-x_{21})+y_{21} && x_{12}<x<x_{22}\\&y=y_2-\sqrt{R_i^2-(x-x_2)^2} && x_{21}<x<x_{22}\\&y=k_3(x-x_{22})+y_{22} && x_{22}<x<x_{12}\end{aligned} \tag{25}$$

We write the equation of motion of the seed for each plot (**Figure 4**)

$$\left.\begin{aligned}m\ddot{x}=mg(\sin a_1-f\cos a_1) && 0<t<t_1\\m\ddot{x}=mg(\sin a_2-f\cos a_2) && t_1<t<t_2\\m\ddot{x}=mg(\sin a_3-f\cos a_3) && t_2<t<t_3\end{aligned}\right\} \tag{26}$$

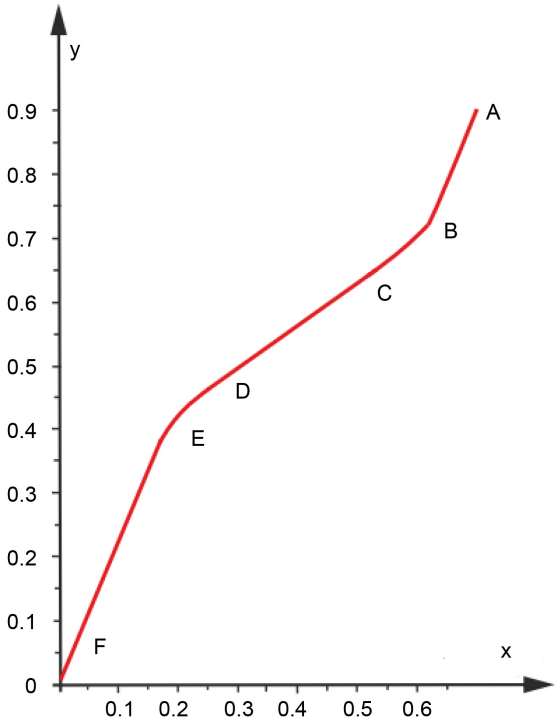


Figure 4. The view of the grate with a concave surface in the plane (x:y).

The differential equations of movement of along the concave profile of the grate, consisting of three broken lines, is integrated on Maple 9.5 under initial conditions, using separate functions, and graphs of the dependence of movement and speed over time are presented (**Figures 5-10**).

Figure 5 and **Figure 6** show the nature of the change in the path and speed of the seeds with time t for various friction coefficients in the aircraft section, for various values of the friction coefficient: 1—f = 0.03; 2—f = 0.06; 3—f = 0.09, in which all graphs are upward.

Figure 7 and **Figure 8** show the nature of the change in the path and speed of passage of seeds with time t at various friction coefficients in the SD section: 1—f = 0.03; 2—f = 0.06; 3—f = 0.09, which also have an upward character of change.

Figure 9 and **Figure 10** show the nature of the change in the speed of seeds along the trajectory s for various friction coefficients in the DE and EF sections for various friction coefficients 1—f = 0.03; 2—f = 0.06; 3—f = 0.09, where the growth rate of parameters in the EF region is much higher than in the previous one.

The graphs show the patterns of change in displacement and speed at different angles, friction coefficient of seeds along grate with a broken line of a concave profile.

The analysis showed the feasibility of using ribs with a concave profile, which helps to accelerate the release of seeds from the working chamber of the fiber separator, which significantly increases the productivity of the cotton processing.

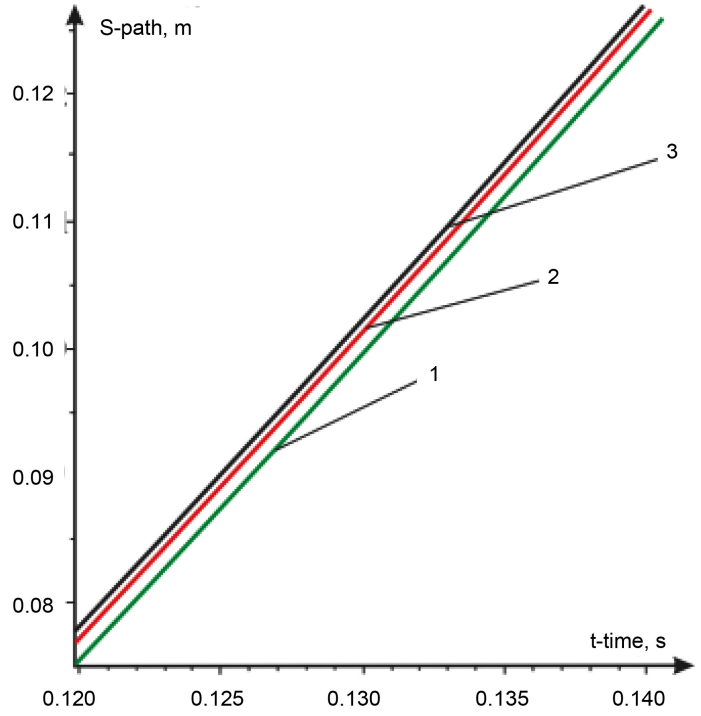


Figure 5. The pattern of change in the path of the seed over time t at various friction coefficients in the plot BC: 1—f = 0.03; 2—f = 0.06; 3—f = 0.09.

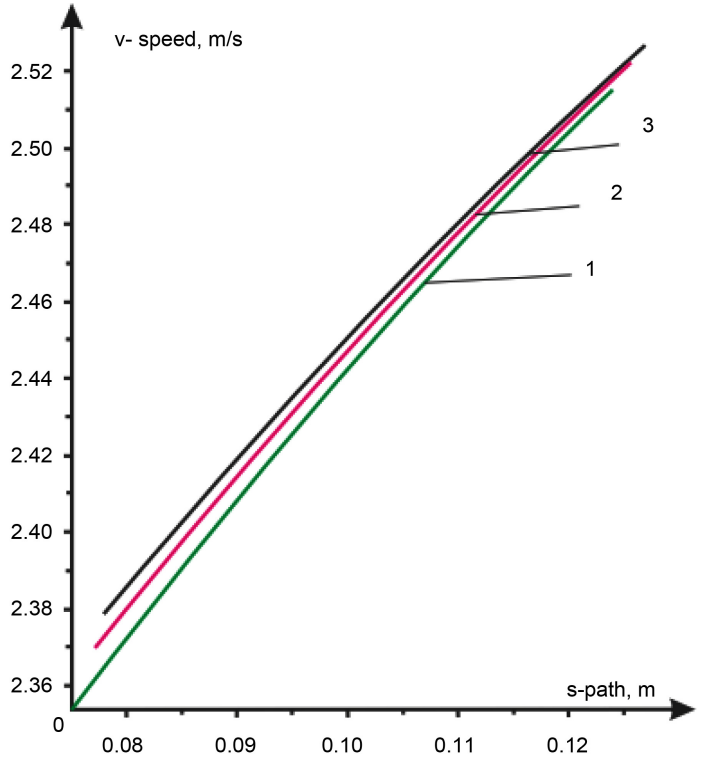


Figure 6. The pattern of change in the path of the seed over time t at various friction coefficients in the plot BC: 1—f = 0.09; 2—f = 0.06; 3—f = 0.03.

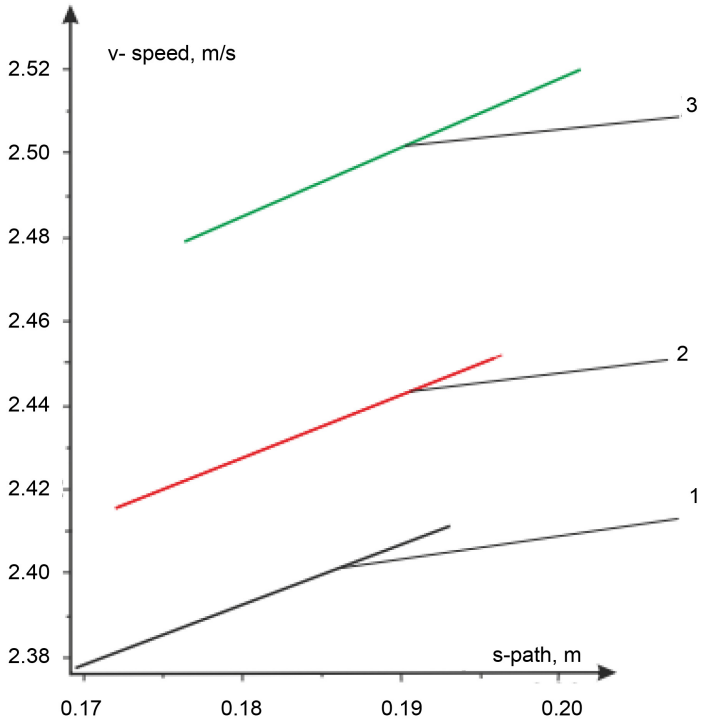


Figure 7. The pattern of change in the path of the seed over time *t* at various friction coefficients in the plot СД: 1—*f* = 0.03; 2—*f* = 0.06; 3—*f* = 0.09.

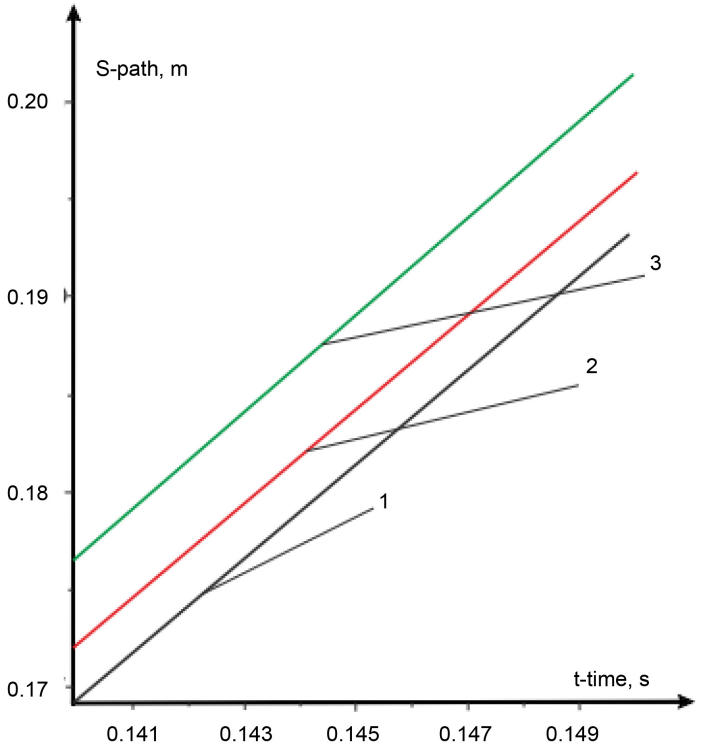


Figure 8. The pattern of change in the path of the seed over time t at various friction coefficients in the plot СД: 1—*f* = 0.03; 2—*f* = 0.06; 3—*f* = 0.09.

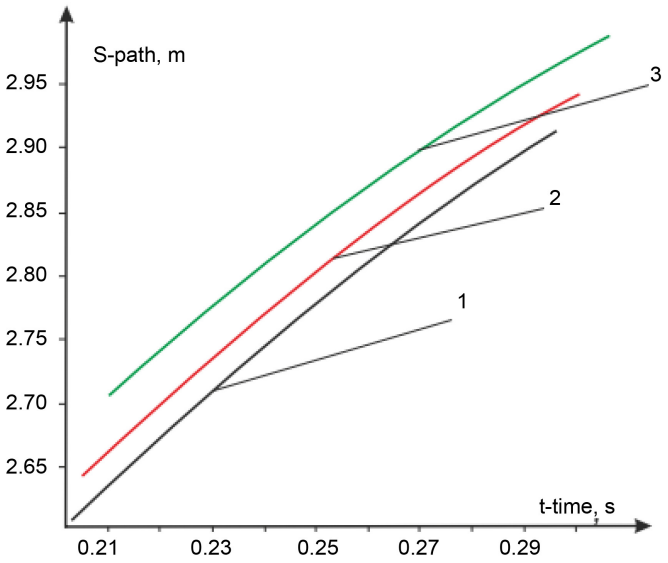


Figure 9. The pattern of change in seed speed along the path s at various friction coefficients in the plot ДЕ: 1—f = 0.03; 2—f = 0.06; 3—f = 0.09.

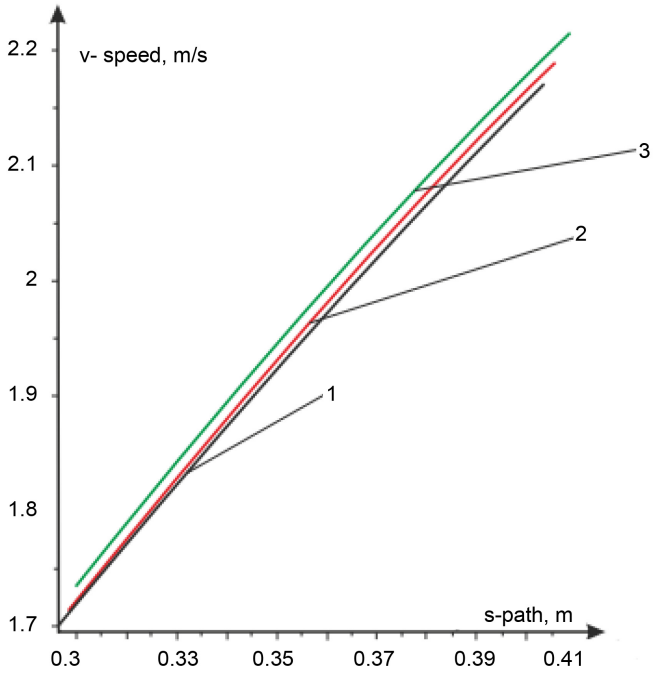


Figure 10. The pattern of change in seed speed along the path s at various friction coefficients in the plot EF: 1—f = 0.03; 2—f = 0.06; 3—f = 0.09.

2. Conclusions

1) An increase in the coefficient of friction leads to a decrease in the movement and speed of the seed in time.

2) You can see that in this section the trajectory and speed of the seeds depend on the friction coefficient, that is, with an increase in the friction coefficient, the

speed decreases.

3) From the above graphs, it is possible to determine the tilt angles $\alpha_1 = 65^\circ, \alpha_2 = 25^\circ, \alpha_3 = 45^\circ$ for the concave grate model with broken, the values of which were accepted, and the grates were tested under production conditions.

4) Determining the movement and speeds of seeds along the grate with a concave profile makes it possible to improve—accelerate the exit of seeds from the roll box of the gin.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Sarimsakov, A., Karimov, A.I. and Murodov, R. (2012) Static Calculation of Processes in the Roll Box of the Saw Gin. *Problems of Mechanics*, No. 2, .17-23.

[2] Akhmedkhodjaev, K.T., Karimov, A.I., Tojiboyev, M. and Sharipov, H. (2013) Determination of the Static Balance of Forces Affecting Seeds Moving along Grates with a Convex Profile. *Scientific and Technical Journal of the Ferghana Polytechnic Institute*, No. 2, 42-45.

[3] Akhmedkhodjaev, K.T., Karimov, A.I., Tojiboyev, M. and Sharipov, H. (2018) Theoretical Analysis of the Movement of Seeds along the Grate with a Convex Profile. *Scientific and Technical Journal of the Ferghana Polytechnic Institute*, No. 3, 35-39.

[4] Karimov, A.I., Azizov, S. and Ismanov, M. (2014) Mathematical Modeling of the Technological Processes Original Processing of Cotton. *International Journal of Innovation and Applied Studies*, **6**, 28-39. http://www.ijias.issr-journals.org

[5] Akramjon, S., Rustam, M., Akmal, U. and Dilmurat, K. (2018) Movement Differential Equation of Seed Roller Which Has Been Installed Stake Accelerator on the Roll Box of Gin Machine. *Engineering*, **10**, 521-529. https://doi.org/10.4236/eng.2018.108038

[6] Akmal, U., Khamit, A., Akramjon, S. and Muazzam, K. (2018) The Saw Gin Stand with Adjustable Movement of the Roll Box. *Engineering*, **10**, 486-494. https://doi.org/10.4236/eng.2018.108034

[7] Khamit, A., Akmal, U. and Kamola, O. (2019) Investigation of the Ginning Process on DП Series Saw Gin Stands. *Engineering*, **11**, 523-530. https://www.scirp.org/journal/eng https://doi.org/10.4236/eng.2019.118036

[8] Obidov, A., Akhmedkhodjaev, Kh., Sarimsakov, O. and Holikov, Q. (2018) Investigation of the Properties of Fibrous Cotton Seeds, for Sorting on a Mesh Surface. *Engineering*, **10**, 572-578. https://doi.org/10.4236/eng.2018.109041

[9] Abdukarimovich, M.O., Ibragimovich, A.K. and Sharipjanovich, S.O. (2018) Designing a New Design of a Loading Cylinder for Pneumo-Mechanical Spinning Machines. *Engineering*, **10**, 345-356. https://doi.org/10.4236/eng.2018.106025

[10] Sarimsakov, O. and Gaybnazarov, E. (2016) About Energy Consumption in Pneumatic Conveying of Raw Cotton. *American Journal of Energy and Power Engineering*, **3**, 26-29.

[11] Sarimsakov, O., Xusanov, C. and Muradov, R. (2016) The Change in Air Pressure

along the Length of the Pipeline Installation for Pneumatic Conveying of Raw Cotton. *Journal of Engineering and Technology*, **3**, 89-92.
http://www.aascit.org/journal/et

[12] Sarimsakov, O. (2016) The Possibility of Reducing Cotton Consumption in Cotton. *American Journal of Science and Technology*, **4**, 68-72.
http://www.aascit.journal/ajst

Call for Papers

Engineering (ENG)

ISSN Print: 1947-3931 ISSN Online: 1947-394X
https://www.scirp.org/journal/eng

Engineering is an international journal dedicated to the latest advancement of engineering. The goal of this journal is to provide a platform for engineers and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of engineering. All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed hard copy. The journal publishes original papers including but not limited to the following fields:

- Aerospace Engineering
- Agricultural Engineering
- Applied Thermal Engineering
- Aquacultural Engineering
- Architecture and Civil Engineering
- Automobile Engineering
- Bioengineering
- Biological Engineering
- Biomedical Engineering
- Biomolecular Engineering
- Bionic Engineering
- Bioprocess and Biosystems Engineering
- Biotechnology and Bioengineering
- Building Services Engineering
- Business & Information Systems Engineering
- Cardiovascular Engineering and Technology
- Cellular and Molecular Bioengineering
- Chemical Engineering & Technology
- Civil Engineering
- Coal Science and Engineering
- Coastal Engineering
- Communication Engineering
- Computational Methods in Engineering
- Computer Aided Engineering
- Computer Applications in Engineering Education
- Computer Engineering
- Computer Methods in Applied Mechanics and Engineering
- Computer-Aided Civil and Infrastructure Engineering
- Crystal Engineering
- Data and Knowledge Engineering
- Earthquake Engineering & Structural Dynamics
- Ecological Engineering
- Electrical and Electronic Engineering
- Energy and Power Engineering
- Energy Engineering
- Energy Systems Engineering
- Engine Engineering
- Engineering Analysis with Boundary Elements
- Engineering and Technology Management
- Engineering Applications of Artificial Intelligence
- Engineering Cybernetics
- Engineering Economy
- Engineering Education
- Engineering Failure Analysis
- Engineering for Sustainability
- Engineering Fracture Mechanics
- Engineering Geology and the Environment
- Engineering in Life Sciences
- Engineering in Medicine
- Engineering Manufacture
- Engineering Materials
- Engineering Mathematics
- Engineering Physics and Thermophysics
- Engineering Science in Medicine
- Engineering Structures
- Engineering Tribology
- Engineering with Computers
- Environmental Engineering
- Environmental Science & Engineering
- Fire Protection Engineering
- Food Engineering
- Food Process Engineering
- Fusion Engineering and Design
- Fuzzy Information and Engineering
- Geophysics and Engineering
- Geotechnical and Geological Engineering
- Green Energy Engineering
- Industrial Engineering
- Landscape and Ecological Engineering
- Macromolecular Materials and Engineering
- Macromolecular Reaction Engineering
- Magnetic Resonance Engineering
- Maintenance Engineering
- Management Engineering
- Maritime Engineering
- Materials Science and Engineering
- Mechanical Engineering
- Mechatronics Engineering
- Medical Engineering & Technology
- Metabolic Engineering
- Metallurgical Engineering
- Microelectronic Engineering
- Micromechanics and Microengineering
- Military Engineering
- Mineral Resources Engineering
- Mining Engineering
- Molecular Engineering
- Nanoengineering
- Natural Gas Science and Engineering
- NeuroEngineering and Rehabilitation
- Nuclear Engineering and Design
- Numerical Methods in Engineering
- Ocean Engineering
- Oil Engineering
- Optics and Lasers in Engineering
- Optimization and Engineering
- Particle Engineering
- Petroleum Engineering
- Petroleum Science and Engineering
- Polymer Engineering and Science
- Power Engineering
- Precision Engineering and Manufacturing
- Probabilistic Engineering Mechanics
- Process Engineering
- Production Engineering
- Protein Engineering, Design and Selection
- Public Safety Information
- Radar Engineering
- Rare Metal Materials and Engineering
- Reliability, Quality and Safety Engineering
- Renewable Energy Engineering
- Requirements Engineering
- RF and Microwave Engineering
- Rock Mechanics and Rock Engineering
- Safety Systems Engineering
- Science and Engineering Ethics
- Social and Environmental Impact of Engineering Innovation
- Software and Knowledge Engineering
- Soil Dynamics and Earthquake Engineering
- Sports Engineering
- Strain Analysis for Engineering Design
- Structural Engineering and Materials
- Surface Engineering and Applied Electrochemistry
- Systems and Control Engineering
- Systems and Software Engineering
- Systems Assurance Engineering and Management
- Test Engineering
- Tissue Engineering and Regenerative Medicine
- Transportation Engineering
- Wind Engineering and Industrial Aerodynamics

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not be previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. For more details, please access the website.

Website and E-Mail

https://www.scirp.org/journal/eng Email: eng@scirp.org

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience

Scientific Research Publishing

Website: https://www.scirp.org

Subscription: sub@scirp.org

Advertisement: service@scirp.org