

# C O M M U N I C A T I O N S

FACULTY OF SCIENCES  
UNIVERSITY OF ANKARA

DE LA FACULTE DES SCIENCES  
DE L'UNIVERSITE D'ANKARA

Series A2-A3: Physical Sciences and Engineering

---

**VOLUME: 66**

**Number: 2**

**YEAR: 2024**

---

Faculty of Sciences, Ankara University  
06100 Beşevler, Ankara –Türkiye  
ISSN 1303-6009 E-ISSN 2618-6462

# C O M M U N I C A T I O N S

FACULTY OF SCIENCES  
UNIVERSITY OF ANKARA

DE LA FACULTE DES SCIENCES  
DE L'UNIVERSITE D'ANKARA

Series A2-A3: Physical Sciences and Engineering

**Volume: 66**

**Number: 2**

**Year: 2024**

**Owner**

Sait HALICIOĞLU, Dean of Faculty of Sciences

**Editor in Chief**

Fatma KARAKOÇ, Ankara University

**Managing Editor**

Şengül KURU, Ankara University

**Area Editors**

İnanç ŞAHİN (Physics) Ankara University, Türkiye	İman ASKERZADE (Askerbeyli) (Computer Eng.) Ankara University, Türkiye
Handan OLĞAR (Engineering Physics) Ankara University, Türkiye	Ziya TELATAR (Electronic Eng.) Başkent University, Türkiye
H. Volkan ŞENAVCI (Astronomy) Ankara University, Türkiye	M. Emin CANDANSAYAR (Geophysical Eng.) Ankara University, Türkiye

**Editors**

Ramiz ALIGULIYEV ANAS, Azerbaijan	Sara CRUZ Y CRUZ SEPI-UPIITA IPN, Mexico	Murat EFE Ankara University, Türkiye
Osman EROGLU TOBB Economy and Tech. Uni., Türkiye	İlhan KOSALAY Ankara University, Türkiye	İsa NAVRUZ Ankara University, Türkiye
Javier NEGRO Univesidad de Valladolid, Spain	Miroslav VOZNAK VŠB – Tech.Uni. of Ostrava, Czech Republic	Emre YENGEL King Abdullah Uni. of Sci.Tech., Saudi Arabia
A. Egemen YILMAZ Ankara University, Türkiye	Kutluay YUCE Ankara University, Türkiye	

**Language Editor**

Fatma KARAKOÇ, Ankara University

**Production Editor**

Şengül KURU, Ankara University

This Journal is published two issues in a year by the Faculty of Sciences, University of Ankara. Articles and any other material published in this journal represent the opinions of the author(s) and should not be construed to reflect the opinions of the Editor(s) and the Publisher(s).

**Correspondence Address:**

COMMUNICATIONS  
EDITORIAL OFFICE  
Ankara University, Faculty of Sciences,  
06100 Beşevler, ANKARA – TÜRKİYE  
Tel: (90) 312-216 87 00 Fax: (90) 312-216 89 00  
e-mail: [commun@science.ankara.edu.tr](mailto:commun@science.ankara.edu.tr)  
<http://communications.science.ankara.edu.tr/index.php?series=A2A3>

**Print:**

Ankara University Press  
İncitaş Sokak No:10 06510 Beşevler  
ANKARA – TÜRKİYE

# C O M M U N I C A T I O N S

FACULTY OF SCIENCES  
UNIVERSITY OF ANKARA

DE LA FACULTE DES SCIENCES  
DE L'UNIVERSITE D'ANKARA

Series A2-A3: Physical Sciences and Engineering

**Volume: 66**

**Number: 2**

**Year: 2024**

## **Research Articles**

---

M.E. YEŞİLYURT, M.S. GÜZEL and E. AKÇAPINAR SEZER, A comparison of machine learning methods for queue length detection .....	132
T. ASUROGLU, Enhancing precision in proton therapy: utilizing machine learning for predicting Bragg curve peak location in cancer treatment .....	140
A. YONAR and Ö. ÖZALTIN, Automatic recognition of coffee bean varieties based on pre-trained CNN architectures.....	162
K. AÇICI, Hemoglobin value prediction with bayesian optimization assisted machine learning models .....	176
I. NAVRUZ, C. DILSIZ, E.S. ORTAK, S.N. BOYRAZ, Optical fiber bending sensor based on speckle pattern imaging.....	201
V. YILMAZ, Sending pictures over radio systems of the trail cam in border security and directing uavs to the right areas.....	214
Y.O. GÜNAYDIN, M. SAHIN and L. AYDIN, Search potential of the high energy-large hadron collider for spin-1/2 excited quarks in di-jet final state.....	228

## A COMPARISON OF MACHINE LEARNING METHODS FOR QUEUE LENGTH DETECTION

Mehmet Eren YEŞİLYURT<sup>1</sup>, Mehmet Serdar GÜZEL<sup>2</sup> and Ebru AKÇAPINAR SEZER<sup>3</sup>

<sup>1</sup>Computer Engineering, Kastamonu University, Kastamonu, TÜRKİYE

<sup>2</sup>Computer Engineering, Ankara University, Ankara, TÜRKİYE

<sup>3</sup>Computer Engineering, Hacettepe University, Ankara, TÜRKİYE

**ABSTRACT.** Queues are formed by people waiting for a service in public institutions and they can be defined as orderly groups of people. Automatically counting the number of people waiting in a queue through video camera footage would provide these institutions with valuable information with regards to customer service quality. In this paper, our goal is to compare several machine learning methods for finding the total number of people waiting in a queue given video camera frames. We approached this problem as a regression task. We used a subset of the Collective Activity Dataset and compared three different methods. The first two methods used bounding box coordinates and orientations provided by the dataset, while the last method utilized the bounding box coordinates to extract feature maps from the frames using RoiAlign. The first method used XGBoost, while the latter methods used Convolutional Neural Networks (CNNs). Results show that the method using RoiAlign presents the best prediction performance in terms of mean squared error and mean absolute error, compared to other methods.



### 1. INTRODUCTION



A queue is a group of people waiting to receive some service in an organized manner. Detecting the number of people waiting in a queue has important applications for retail stores and public institutions such as hospitals. Knowing the queue length density for different time periods would allow public-serving facilities to optimize their human resource allocation in order to reduce waiting times and improve their quality of service.

Compared to a crowd, a queue is characterized by its orderly structure. The people comprising a queue form straight or curved lines, and they usually face the same

*Keywords:* Queue length detection, crowd counting, machine learning.

 myesilyurt@kastamonu.edu.tr-Corresponding author;  0000-0002-7322-5572

 mguzel@ankara.edu.tr;  0000-0002-3408-0083

 ebru@hacettepe.edu.tr;  0000-0002-9287-2679.

general direction. The known literature mostly focuses on the crowd counting problem, which has a similar objective to the queue length detection problem. Crowd counting can be implemented through detection, regression or density map estimation [7] methods. Detection based methods are used to detect hand-crafted features such as the existence of body parts or body appearance [6] to find the total body count. Object detection techniques such as YOLO and Faster-RCNN have also been used for crowd counting [8, 9, 10, 11]. In contrast to object detection methods, regression based methods predict the total count of people in an image directly. Finally, the density map estimation methods work by creating a crowd density map which is then used to find the total count. Since regression and density map estimation methods do not rely on detecting and counting people individually, they are more resilient to occlusion problems.



FIGURE 1. A sample frame from the collective activity dataset [1].

ACTi, which is a corporation that provides video analytics solutions, offers a queue management system that can determine the amount of time a person has spent waiting on a queue [13]. The system works by counting the number of people in a region of interest determined by store managers [14].

Saini et al. used bounding boxes obtained from an object detector to estimate the number of people waiting in the queue in a given frame [15]. Their method assumes that a queue lies on a straight line on the image and fits a line in the form of  $y=mx+b$

that minimizes the distance to the midpoints of the bounding boxes. To estimate the parameters of this line, they formulate the problem as an optimization problem and solve it using convex hull optimization.

In [16], Wu et al. model scenes as `actor-relation graphs` in order to classify collective and individual activities carried out by people in video streams. They use RoiAlign to extract features for each bounding box, which is then used to obtain feature vectors. These vectors are then used to build actor-relation graphs. After graph convolution and pooling, classifiers predict group activity and individual activity carried out by actors.

In this paper, we approached the queue length detection task as a regression problem. Given an image, our objective was to count the total number of people waiting in a queue while disregarding those that were not part of a queue. To this end, we evaluated 3 different methods on a subset of the collective activity dataset.

## 2. DATA AND METHODS

The Collective Activity dataset was created by Choi et al. with the goal of classifying collective activities carried out by groups of individuals [1]. The dataset is made up of 44 short videos. In each video, there is a small number of people carrying out a group activity such as talking, queuing or walking. Labels provided for each video contain the bounding box coordinates, individual activity and pose for every tenth frame.

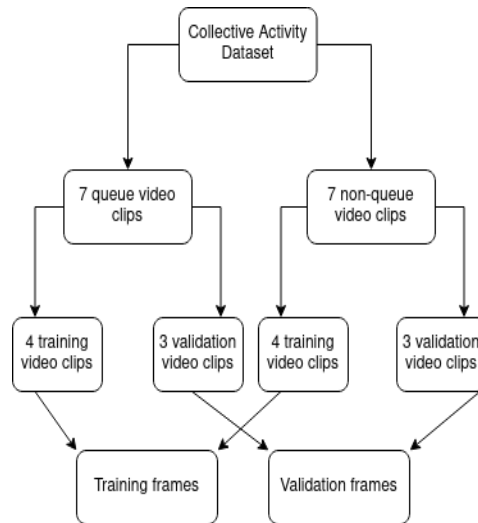


FIGURE 2. Dataset preparation steps.

To obtain the subset of the Collective Activity dataset used in this study, we used all 7 of the video clips that contained queues as well as 7 video clips that contained no queues. For both video sets, 4 out of 7 video clips were chosen for the training and the remaining 3 video clips were reserved for validation. Separating video clips in this manner ensured that the frames in the validation set and the frames in the training set came from different video clips. From each video clip in the training set, 21 frames were selected equidistantly. Similarly, 7 frames were extracted from every video clip in the validation set. This process resulted with 168 frames in the training set and 42 frames in the validation set with a validation/train split ratio of 20%. For each frame, the total number of individuals labeled as `queuing` served as the target variable of the regression task.

The first method we used was XGBoost utilizing tabular features. XGBoost, which stands for `Extreme Gradient Boosting`, was created by Chen and Guestring and it is defined as a `scalable tree boosting system` [2]. We used midpoint coordinates of each bounding box and pose information as features. Since each frame contains a variable number of individuals, and therefore, a variable number of bounding boxes; there are a different number of features for each frame. To get a fixed number of features to be given to the model, we assumed a limit of 14 people in a single frame. If the number of people in the frame was less than this value, the features of the individuals that were not present were set to 0. We used the same limit in the other methods as well. For the XGBoost model, we ended up with 42 features for each frame: 14 X coordinates, 14 Y coordinates and 14 pose labels.

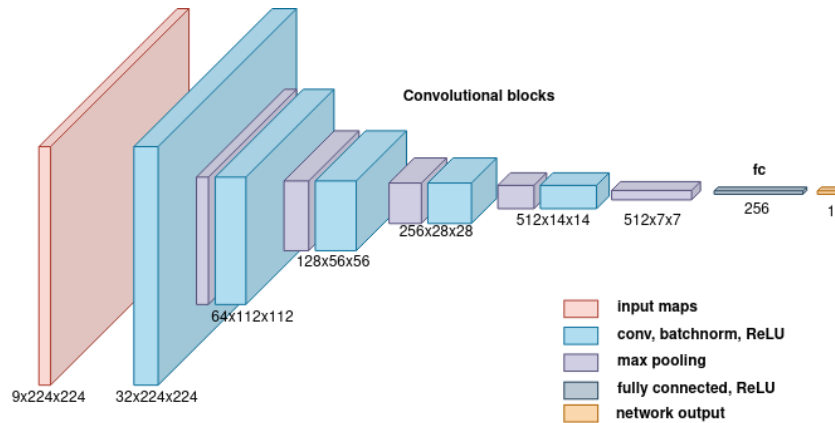


FIGURE 3. CNN model architecture.

The second method utilized a convolutional neural network (CNN). CNNs are deep learning models that are appropriate for input data that has regular spatial

structure, such as images [3]. The CNN architecture we used can be seen in Figure 3. Similar to VGG, the architecture is comprised of a number of convolutional layers with  $3 \times 3$  kernels and max pooling, followed by a fully-connected layer. The input dimensions are  $9 \times 224 \times 22$ . The channels in the input correspond to different human orientations in the dataset. To create the input for the model, we obtained the middle point coordinates of each bounding box in a frame and scaled these coordinates to fit in a  $224 \times 224$  grid. We then marked these coordinates in their respective pose channel. For example, if a person's rescaled bounding box coordinates were calculated as (100, 200) and that person's orientation was labeled as facing right, then the respective coordinate in the first channel (which corresponds to the orientation `right`) was set to 1. The persons with unspecified orientations were placed in the first channel.

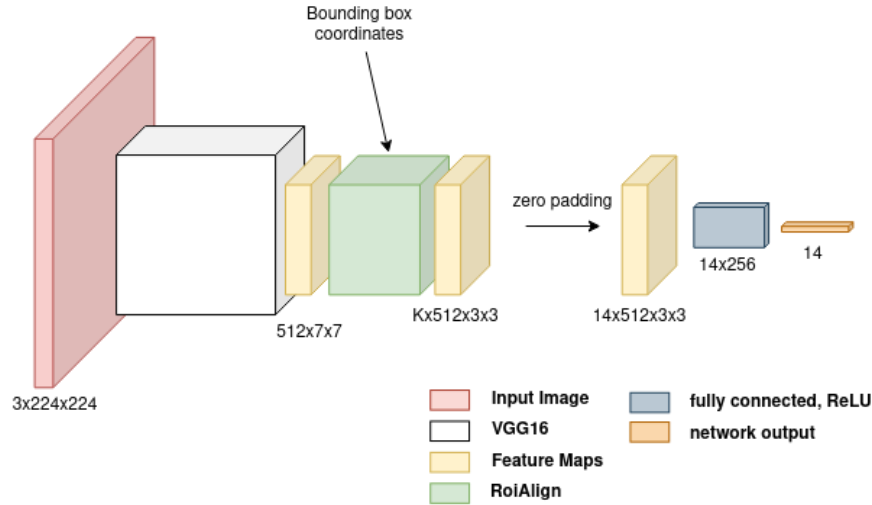


FIGURE 4. RoiAlign model architecture.

The third method also makes use of convolutional neural networks and its architecture can be seen in Figure 4. This network corresponds to the feature extractor stage in [16] by Wu et al. The model is made up of a feature extractor followed by a classifier that utilizes RoiAlign, which was introduced in [4] by He et al for extracting feature maps from a region of interest. The feature extractor takes a  $3 \times 224 \times 224$  input image and creates feature maps of size  $512 \times 7 \times 7$ . We used a VGG16 [5] model without the fully-connected layers as the feature extractor. The classifier takes the output of the feature extractor and bounding box positions as input. Using RoiAlign, the classifier extracts  $3 \times 3$  feature maps for each bounding box in the frame. This is followed by fully-connected layers. The model outputs a



value for each bounding box in the frame that represents the probability of the person in the bounding box belonging to a queue. We can then use the number of positive predictions output by the model as the number of people predicted to be in a queue for a given frame.

The XGBoost model was trained with default hyperparameters. To train the CNN and CNN RoiAlign models, we used the Adam [12] optimizer. After trying learning rates that ranged from  $10^{-5}$  to  $10^{-2}$ , we used a learning rate of  $10^{-4}$  for the CNN model and a learning rate of  $10^{-5}$  for the CNN RoiAlign model. Each training run lasted for 50 epochs, with early stopping after 10 epochs with no improvement.

TABLE 1. Evaluation results (Average of 5 training runs).

Method	Mean squared error	Mean absolute error
XGBoost	11.20	2.22
CNN	10.58	2.99
CNN RoiAlign	4.28	0.97

### 3. RESULTS

The results are shown on Table 1. We used mean squared error (MSE) and mean absolute error (MAE) to evaluate and compare each method. Due to the small size of the test set, there is a possibility of high variance in the obtained results. Because of this, we trained and evaluated each method 5 times and presented the average values obtained from these training runs. The network that utilized RoiAlign had the best MSE and MAE scores out of all methods. The CNN model that only used bounding box coordinates and pose information was very similar to the XGBoost model in terms of the MSE, while the XGBoost model gave better results in terms of the MAE.

According to the results shown in Table 1, it can be seen that the CNN RoiAlign method yields the best results compared to other methods, while the results for XGBoost and CNN in terms of the mean squared error were similar. The RoiAlign method makes use of extracted feature maps from the input image, while the XGBoost and the CNN models used only the bounding box position and pose information. The lower performance of the XGBoost and CNN models may have been caused by their direct dependency to bounding box coordinates provided in pixel coordinates. Since the dataset contained a mix of indoor and outdoor scenes with different locations and camera angles, the physical distances related to pixel distances were different for each scene.

#### 4. CONCLUSION

This paper presents a comparison of different Machine and Deep Learning based methods for Queue Length Detection problem. In order to conduct experiments, a subset of the Collective Activity Dataset was employed. Three different methods were compared. While the first two methods, XGBoost and CNNs, used bounding box coordinates and orientations provided by the dataset; the final method, based on CNNs architecture, utilized the bounding box coordinates to extract feature maps from the frames using RoiAlign. Results showed the superiority of the final method over previous methods. It should be noted that the lower performance of the first two methods may have been caused by their direct dependency on pixel coordinates. Even though the CNN RoiAlign method still utilizes pixel coordinates of the bounding box locations, we found that using RoiAlign to extract feature maps for each bounding box and then classifying these boxes individually gives better results compared to other methods. Authors are planning to apply more complex Deep Learning architectures and larger datasets to the problem at hand in future work.

**Author Contribution Statements** Mehmet Eren Yeşilyurt analyzed the data, performed the computation work, prepared figures and/or tables. Mehmet Serdar Güzel conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work. Ebru Akçapınar Sezer analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

**Declaration of Competing Interests** The authors declare that they have no competing interests.

**Acknowledgement** This study is supported by TUBITAK TEYDEB (Project Name: EYEIN: HASTANE İÇİ ORTAM İZLEME VE DURUM TESPİTİ İLE KALABALIK YÖNETİMİ, Project No: 3220818).

#### REFERENCES

- [1] Wongun, C., Shahid, K. and Savarese, S., What are they doing?: Collective activity classification using spatio-temporal relationship among people, *IEEE 12th Inter. Conf. on Comp. Vis. Wor. ICCV Work.*, (2009), 1282-1289, <https://doi.org/10.1109/ICCVW.2009.5457461>.
- [2] Chen, T. and Guestrin, C., XGBoost, *Proc. of the 22nd ACM SIGKDD Int. Conf. on Know. Disc. and Data Min.*, (2016).
- [3] Yamashita, R., Nishio, M., Do, R. K. G. et al., Convolutional neural networks: an overview and application in radiology, *Ins. Ima.*, 9 (2018), 611-629, <https://doi.org/>


- 10.1007/s13244-018-0639-9.
- [4] He, K., Gkioxari, G., Dollar, P. and Girshick, R., Mask R-CNN, *arXiv:1703.06870*, (2018), <https://doi.org/10.48550/arXiv.1703.06870>.
  - [5] Simonyan, K. and Zisserman, A., Very deep convolutional networks for large-scale image recognition, *arXiv:1409.1556*, (2015), <https://doi.org/10.48550/arXiv.1409.1556>.
  - [6] Khan, M. A., Menouar, H. and Hamila, R., Revisiting crowd counting: State-of-the-art, trends, and future perspectives, *Imag. and Vis. Comp.*, 129 (2023), 104597, <https://doi.org/10.1016/j.imavis.2022.104597>.
  - [7] Sindagi, V. A. and Patel, V. M., A survey of recent advances in CNN-based single image crowd counting and density estimation, *Pat. Rec. Let.*, 107 (2018), 3-16, <https://doi.org/10.1016/j.patrec.2017.07.007>.
  - [8] Ruchika, R., Purwar, K. and Verma, S., Analytical study of YOLO and its various versions in crowd counting, *Int. Data Com. Tech. and Int. of Thi.*, (2022), 975-989, [https://doi.org/10.1007/978-981-16-7610-9\\_71](https://doi.org/10.1007/978-981-16-7610-9_71).
  - [9] Valencia, I. J. C., Dadios, E. P., Fillone, A. M., Puno, J. C. V., Baldovino, R. G. and Billones, R. K. C., Vision-based crowd counting and social distancing monitoring using tiny-YOLOv4 and DeepSORT, *IEEE Int. Sma. Cit. Conf. (ISC2)*, (2021), 1-7, <https://doi.org/10.1109/ISC253183.2021.9562868>.
  - [10] Muzamal, J. H., Tariq, Z. and Khan, U. G., Crowd counting with respect to age and gender by using faster R-CNN based detection, *Int. Conf. on Appl. and Eng. Math. (ICAEM)*, (2019), 157-161, <https://doi.org/10.1109/ICAEM.2019.8853723>.
  - [11] Akbar, N. and Jamal, E. C., Crowd counting using region convolutional neural networks, *8th Int. Conf. on Elec. Eng., Comp. Sci. and Info. (EECSI)*, (2021), 359-364, <https://doi.org/10.23919/EECSI53397.2021.9624288>.
  - [12] Kingma, D. P. and Ba, J., Adam: A method for stochastic optimization, *arXiv:1412.6980*, (2017), <https://doi.org/10.48550/arXiv.1412.6980>.
  - [13] People Queue Detection - ACTi Corporation. Available at: <https://www.acti.com/technologies/people-queue-detection>. [Accessed July 2023].
  - [14] ACTi Corporation, Queue management whitepaper, ACTi Corporation, (2023). Available at: <https://download.acti.com/?id=10618>.
  - [15] ShivamJalotra, Queue-Detection, (2023). Available at: <https://github.com/jalotra/Queue-Detection>. [Accessed July 2023].
  - [16] Wu, J., Wang, L., Wang, L., Guo, J. and Wu, G., Learning actor relation graphs for group activity recognition, *arXiv:1904.10117*, (2019), <https://doi.org/10.48550/arXiv.1904.10117>.

## ENHANCING PRECISION IN PROTON THERAPY: UTILIZING MACHINE LEARNING FOR PREDICTING BRAGG CURVE PEAK LOCATION IN CANCER TREATMENT

Tunc ASUROGLU<sup>1</sup>

<sup>1</sup>Faculty of Medicine and Health Technology, Tampere University,  
33720 Tampere, FINLAND

**ABSTRACT.** In proton beam therapy, the Bragg peak is the point where protons lose energy the fastest. This point is crucial for dose control, preserving healthy tissues, minimizing lateral scattering, and the success of treatment planning. However, accurately predicting the location of the Bragg peak is challenging due to the complex interactions of protons with tissues. This study proposes a machine learning (ML) approach to predict the exact location of the Bragg peak from phantom tissue proton beam therapy experiments. A dataset comprising the eight most commonly used biomaterials, which mimic human tissue in proton therapy procedures, has been curated for this study. Various ML models are benchmarked to find the most successful approach. ML model parameters are further optimized using a metaheuristic approach to achieve the highest prediction capability. In addition, feature contributions of each feature in the dataset are analyzed using an explainable artificial intelligence (XAI) technique. According to experimental results, Random Forest (RF) model that is optimized with Genetic Algorithm (GA) achieved 0.742 Correlation Coefficient (CC) value, 0.069 Mean Absolute Error (MAE) and 0.145 Root Mean Square Error (RMSE) outperforming other ML models. The proposed approach can track and predict the movement of the proton beam in real-time during treatment, enhancing treatment safety and contributing to the more effective management of the treatment process. This study is the first to predict exact Bragg curve peak locations from proton beam therapy experiments using ML approaches. The optimized ML model can provide higher precision in identifying the needed beam dosage for targeted tumor and improving treatment outcomes.

*Keywords:* Machine learning, regression, proton therapy, Bragg curve.  
✉ tunc.asuroglu@tuni.fi-Corresponding author;  0000-0003-4153-0764.

## 1. INTRODUCTION

Protons are positively charged and heavy, so they lose energy slowly but continuously as they scatter through matter. This scattering is at small angles [1, 2, 3]. The Bragg peak is the point where protons lose energy the fastest [4, 5]. Because tumors are large, a single-energy proton beam is not enough to treat them. Instead, proton beams with different energies are needed [4, 5, 6, 7]. It is important that the target is made of a material that is similar to tissue [8]. This is because the dose of radiation delivered to the target should be as accurate as possible. Phantoms are used to simulate the target and help optimize the dose [1, 8]. This process continues until all energies are depleted, and then they suddenly come to a halt. The dose accumulation process forms the characteristic depth-dose curve ("Bragg curve") of a broad monoenergetic proton beam. The highest dose point is called the Bragg peak. The depth of the peak, i.e., the range of protons, depends on the initial energy. Detecting the location of this peak is crucial for dose control, preserving healthy tissues, minimizing lateral scattering, and the success of treatment planning [7].

It is generally accepted that the accuracy of the results of a proton beam therapy simulation is related to the similarity between the phantom material and the tissue it is simulating [8]. The International Atomic Energy Agency (IAEA) recommends using water as the phantom material for soft tissue simulations, because it is easy to obtain and has a density similar to soft tissue [9, 10, 11]. Other biomaterials, such as those with mass densities similar to hard tissue, can also be used [12, 13, 14, 15]. The dose delivered to the target, the shape of the Bragg peak, and the results of nuclear interactions are all crucial factors to consider when evaluating the properties of a phantom material [8, 11]. Even though they are less commonly used, biomaterials are still important for simulating interactions such as backscattering, collision events, phonon production, and side scattering [1, 12].

The machine learning (ML) model, trained with information that are originated from proton beam simulations using tissue like biomaterials, can help researchers better understand the interactions of protons within tissues. ML approaches can provide higher precision in identifying the needed beam dosage for targeted tumor during treatment planning and achieving dose optimization. Accurate detection of the Bragg peak of the proton beam allows maximum focus of the treatment dose on the tumor region and enables finding methods to minimize damage to normal tissues. Additionally, the ML model can track the movement of the proton beam in real-time during treatment and quickly provide alerts in case of any deviations. This can enhance treatment safety and contributes to the more effective management of the treatment process.

In the domain of proton beam therapy, where precision and effectiveness are paramount, the importance of optimizing ML models and conducting feature impact

analysis cannot be overstated. Proton therapy, renowned for its capacity to precisely target tumors while preserving healthy tissues, stands to gain immensely from the application of ML [16, 17, 18]. The optimization of ML models becomes instrumental in refining treatment parameters, ensuring accurate predictions of pivotal factors such as the Bragg peak curve location. Through the utilization of sophisticated algorithms, clinicians can tailor ML models to accommodate diverse patient profiles, optimizing treatment plans to achieve maximum therapeutic impact. Feature impact analysis also plays a pivotal role by unveiling the influential variables that significantly impact treatment outcomes when using the proton beam therapy [19]. This insight into feature impacts enables personalized adjustments in proton beam therapy, contributing to more tailored and efficacious cancer treatments. The seamless integration of ML model optimization and feature impact analysis not only elevates the precision of proton beam therapy but also represents a substantial leap toward the implementation of individualized and optimized strategies in cancer care.

To this end, a dataset is constructed with most commonly used biomaterials, which mimics human tissue in proton therapy procedures. Afterwards various ML models are benchmarked to find out which one is better on finding the exact location of Bragg curve peaks under different energy levels and with different biomaterials. To achieve a robust model, ML model parameters are further optimized by using genetic algorithm optimization method. As a last step, feature contributions are assessed by SHAP (SHapley Additive exPlanations) technique to see which features are important when making the predictions.

Contributions of this study can be summarized as follows:

- This is the first study that aims to predict exact Bragg curve peak locations from proton beam therapy experiments using ML approaches.
- A dataset comprising the eight most commonly used biomaterials, which mimic human tissue in proton therapy procedures, has been curated for this study.
- Various ML models are benchmarked to find out which one is better on finding the exact location of Bragg curve peaks under different energy levels and with different biomaterials.
- ML model parameters are further optimized using a metaheuristic approach to achieve the highest prediction capability.
- Feature contributions of each feature in the dataset are analyzed using an explainable artificial intelligence (XAI) technique.
- By using ML approaches that are optimized by metaheuristic algorithms, this approach can track and predict the movement of the proton beam in real-time during treatment. This can enhance treatment safety and contributes to the more effective management of the treatment process.

The paper is organized in the following manner: Section 2 presents details of the proposed framework, covering the dataset, feature extraction process, machine

learning approaches, genetic optimization algorithm and feature importance assessment. Section 3 explains the evaluation metrics and highlights the experimental results. Ultimately, Section 4 concludes the paper with conclusion and discussion section.

## 2. MATERIAL AND METHODS

**2.1. Proposed machine learning framework.** Proposed ML framework leverages biomaterial features and energy levels as input and employ a ML model for establishing relationships between biomaterials and Bragg curve peak points. The framework involves a straightforward learning process, which encompasses training and testing/evaluation stages. To begin, features are extracted from particle therapy experiments, a ML model is trained using these feature vectors, where each sample corresponds to a specific peak location value. In the last step, the model predicts the peak point of a test sample. The trained model's performance is assessed using various evaluation metrics. The workflow of proposed approach is illustrated in Figure 1.

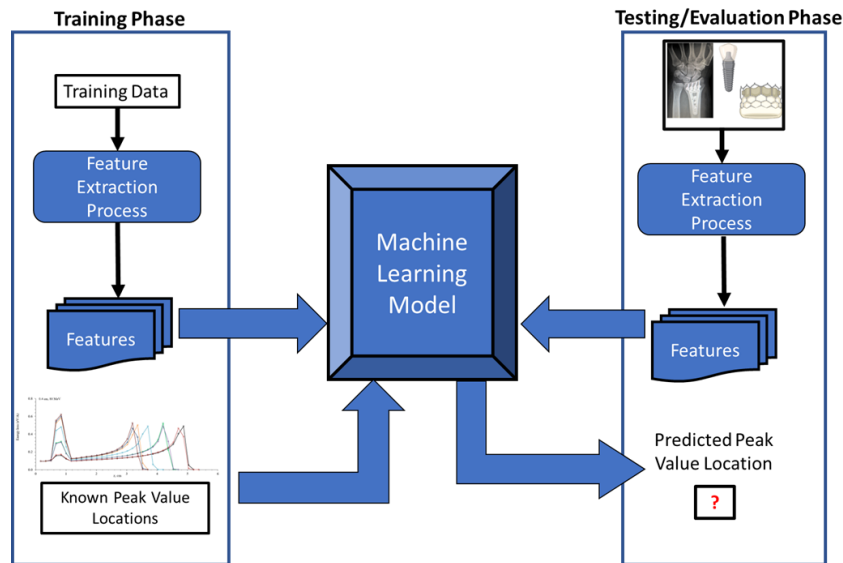


FIGURE 1. Workflow of proposed approach.

**2.2. Dataset and Feature Extraction.** In this study, the necessary data for ML models were obtained from the Monte Carlo (MC) Transport of Ions in Matter

(TRIM) simulation tool. This tool can calculate ion interactions within the target [20]. Information such as the type and energy of ions, selected phantom type and shape, incident angle of the beam, parameters to be calculated, particle count, and probability can be inputted in TRIM [20]. TRIM can calculate all kinetic events related to the energy loss processes of ions, damage inflicted on the target, scattering, ionization, voids in the crystal structure of polymeric biomaterials, phonon generation, and recoil [20]. All target atom cascades (polymeric, soft tissue, and water) in the selected phantom can be tracked and recorded in detail [21]. The recorded data aims to determine the most suitable phantom biomaterial and obtain biomaterials that are more similar to human tissue. MC TRIM and feature extraction procedure is demonstrated in Figure 2. First, experiments are conducted using MC TRIM algorithm that includes proton beam reflected on biomaterial and as a second stage biomaterial features, and Bragg curve peak locations are recorded for each experiment.

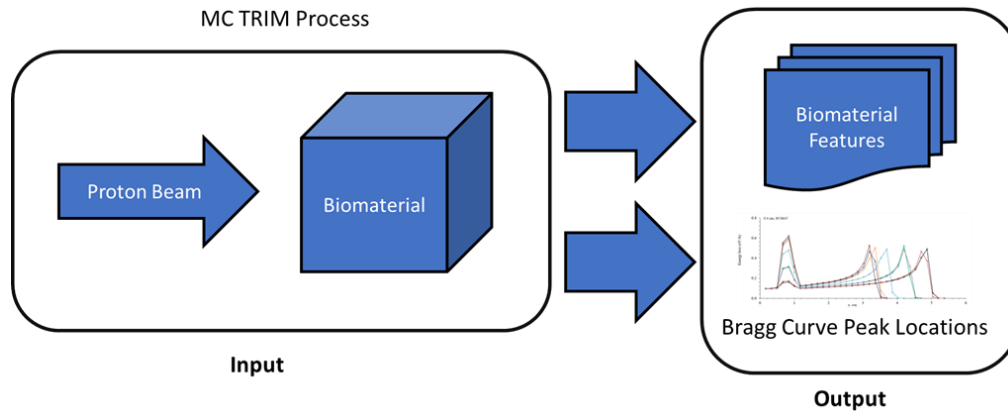


FIGURE 2. Experimental procedure to obtain Bragg curve peak locations and biomaterial features.

The experiments are conducted on 8 biomaterials: Cortical bone, Teflon, Titanium alloy, Aluminum oxide, stainless steel, Vitalium, cobalt-nickel-chromium-molybdenum, and Nital. Each biomaterial is tested with five thicknesses (0.4, 0.6, 0.8, 1, and 1.2 centimeters (cm)) and 10 different energy levels (80, 100, 120, 140, 160, 180, 200, 220, 240, and 250 Mega-electronvolt (MeV)). This combination results in a dataset of 400 samples for the machine learning models to perform predictions. Each sample in the dataset has a feature set containing information relevant to predicting the Bragg curve peak point location, which ranges between 0.23 and 1.43 cm. These features are:

- *Energy in MeV*: Energy value of the proton.



- *Biomaterial thickness (cm)*: This directly affects the depth of the Bragg peak within the material.
- *Biomaterial mass density*: This property influences how much radiation the material interacts with.
- *Biomaterial atomic density*: Similar to mass density but considers the number of atoms per unit volume.
- *Biomaterial atomic composition percentage of each atom*: This captures the elemental setup of the material, which impacts its interaction with radiation. To create this feature, a feature vector is constructed for each unique atom present in the dataset. Each biomaterial sample's corresponding feature vector is filled with its respective atomic percentages. Any missing elements (not present in the specific biomaterial) are represented by zeros in the vector.

To ensure all features contribute equally to the machine learning models, the dataset undergoes normalization. This process centers the mean value of each feature to zero and scales the standard deviation to one. This standardization prevents features with larger scales from dominating the model's learning process.

**2.3. Selected Machine Learning Models for Proposed Framework.** In order to assess the effectiveness of ML models on predicting the exact peak location of the Bragg curve, various models are evaluated. Since the prediction task in this study is a regression problem due to predicting exact values instead of class labels, models that utilize regression process are selected. These models are Decision Tree (DT), Random Forest (RF), Linear Regression (LR), eXtreme Gradient Boosting (XGBoost), Support Vector Regression (SVR) and k-nearest neighbor (kNN).

The Decision Tree (DT) algorithm can manage both numerical and categorical data, seeking the feature that best divides the training set [22]. This feature is selected based on its maximum information gain. Upon evaluating the potential values of this feature, the algorithm branches into sub-trees and assigns target values. Meanwhile, it explores other features with high information gain. This iterative process continues until a clear decision is made regarding the combination of features that forms a definitive rule for predicting target values. By the end of the algorithm, all features have been assessed, and every sample has been assigned an appropriate target value [22]. Its simplicity and reliability have made it a popular ML tool across various domains [23, 24].

Random Forest (RF), a member of the decision tree family, employs an ensemble learning technique to enhance its predictive power [25]. This algorithm has gained immense popularity for its ability to effectively combat overfitting in both classification and regression tasks, while maintaining relatively low computational demands [25, 26, 27]. RF builds a multitude of decision trees by randomly sampling subsets of data, known as bootstrap samples [25]. Unlike traditional decision tree algorithms that strive to identify the optimal variable at each decision point, RF

introduces an element of randomness by considering a random subset of variables at each split. This approach is primarily implemented to mitigate the correlation among individual decision trees [25]. Such correlation can negatively impact predictions and hinder overall performance. The incorporation of randomness in RF is crucial for sound decision-making. Highly correlated variables can lead to biased predictions and suboptimal outcomes [26]. By introducing randomness, RF effectively reduces the influence of individual variables, allowing the algorithm to make more robust and unbiased decisions. The predictions from these independent decision trees are then aggregated to achieve the final outcome [26]. This ensemble approach not only addresses overfitting but also enhances the overall accuracy of the algorithm.

Linear regression (LR) stands out as one of the most fundamental and widely employed regression techniques, known for its simplicity [28]. One of its key advantages lies in the straightforward interpretability of its outcomes. In essence, linear regression fits a linear model characterized by coefficients to minimize the residual sum of squares between the observed target values in the dataset and those predicted by the linear model. Using this model, it becomes possible to make predictions for unknown target values by utilizing specified parameters along with the computed coefficients [29].

eXtreme Gradient Boosting (XGBoost) is an ensemble ML approach build on decision trees, utilizing an iterative function called gradient descent framework [30]. This iterative approach continually enhances model performance by enhancing the learning capacity of weak learners. XGBoost is versatile and applicable to both classification and regression problems, making it a valuable tool for addressing supervised learning tasks. Ensemble learning, a fundamental concept, involves creating multiple weak predictors to make predictions for a dataset and subsequently combining these individual predictions using a specific strategy to arrive at the final prediction result [31]. XGBoost represents an advancement over the traditional gradient boosting decision tree algorithm, offering improvements in terms of model building speed, prediction ability, and adjustability. In contrast to gradient boosting, XGBoost incorporates regularization within the loss function to formulate its objective function [32].

Support Vector Regression (SVR) represents a crucial facet of the broader Support Vector Machine (SVM) framework [33]. Unlike SVM classification, where multiple classes of sample points are involved, SVR is specifically tailored for situations where only one type of sample point is present. The fundamental objective of SVR differs from SVM as it does not seek to maximize the margin or separation distance between multiple types of sample points. Instead, SVR's goal is to minimize the collective deviation between the sample points and a hyperplane [34]. In the case of addressing nonlinear problems, SVR leverages a kernel function to transform the

nonlinear regression task into a higher-dimensional space. This transformation allows SVR to identify an optimal hyperplane for effectively separating the sample points in this transformed space, thereby facilitating accurate regression in cases where linear relationships may not hold [35].

The k-Nearest Neighbors (kNN) model is a supervised ML technique primarily employed for classification purposes [36]. This algorithm revolves around a flexible parameter, denoted as 'k' which represents the number of nearest neighbor's to consider when doing predictions. The kNN algorithm operates by identifying the nearest samples or neighbors within a training dataset in response to a query sample. These nearest samples are determined based on their proximity to the query sample. Once the k nearest samples are identified, the algorithm employs a majority voting rule to determine which class appears most frequently among them. The class with the highest frequency is designated as the final classification for the given query [37]. For regression problems as in this study, weighted average of the prediction value is calculated, where closer neighbors have more influence on the prediction.

**2.4. Genetic Algorithm (GA) for Machine Learning Model Parameters Estimation.** When building ML models, finding, and estimating the model parameters can be a crucial task. Improper parameter sets can lead to weak predictors. To solve this problem there are several ways, a classical approach, the grid-search is a method for parameter optimization in which a predefined range of parameter combinations is exhaustively evaluated to identify the most effective configuration for a model. Each combination in the grid is assessed for model performance using a chosen scoring metric. The set of parameters that produces the best model performance is usually chosen as the optimal configuration [38]. But this process can advance slowly and obtaining an optimal parameter set can be difficult [39]. Heuristic methods such as genetic algorithm and evolutionary algorithms can be used to find approximate solutions to computationally expensive problems more quickly and efficiently than conventional methods [40]. These methods have been used to solve numerical problems and prediction problems [41, 42]. Heuristic methods aim for feasible solutions within the problem domain. Starting from a candidate solution, they produce a new generation of solutions with modified objective values.

In this study, in order to optimize the parameters of ML models, genetic algorithm is used. In this study, the best model is selected according to the achieved results and then as a second step, GA optimization is applied to this model for further optimizing the parameter set for better performance.

GA is a type of heuristic search algorithm that mimics the process of natural selection to find optimal solutions to complex problems. It is considered as a probabilistic optimization method because GA uses randomness to explore different solutions and find the best one [43]. When employing GA, every potential solution

is represented as a chromosome within the problem's search space. Search space corresponds to the population. From a biological perspective, these chromosomes mirror the traits of an organism. The distinguishing factor is that the genes contained within different chromosomes exhibit variations. Genes are encoded using a binary encoding technique (0 or 1). Basic units of operations in GA are the chromosomes, three operators adjust the chromosomes to achieve the optimal solution. These operators are selection, crossover, and mutation operator [44].

- *Selection operator*: Throughout the course of biological evolution, all living beings must undergo adaptation to their surrounding environment, and it's only those individuals who successfully navigate and align with the demands of their habitat that emerge as superior individuals. This phenomenon is recognized as natural selection, and the selection mechanism in GA emulates this natural process. A chromosome's likelihood of elimination diminishes as its fitness score increase.

- *Crossover operator*: This operator is the most crucial one in the algorithm. It entails the exchange of gene segments from two different chromosomes, resulting in the creation of two completely new chromosomes.

- *Mutation operator*: In a biological perspective not every chromosome exchanged, some of it actually mutates to generate previously unseen new chromosomes. This probability of mutation actually helps prevent the algorithm from prematurely converging to an undesired state. It ensures that evolution is more varied and enhance the GA's capacity for local search.

In the context of GA parameter optimization, chromosome denotes the parameter values to be optimized, while the search space relates to the parameter boundaries. Within the population, chromosomes which are made of genes, represent the parameters that require optimization for the desired problem. The objective of the GA is to seek the optimal individual that meets the criteria of the fitness function. Overall working mechanism of optimization is given in Figure 3.

**2.5. Feature importance assessment using SHAP (SHapley Additive exPlanations) values.** In this study, one of the explainable artificial Intelligence (XAI) technique called SHAP is utilized to evaluate and determine the feature importance on predicting the peak point location of the Bragg curve [45]. The SHAP method is used to compute SHAP values for individual features within a ML model, enabling a better grasp of how these features impact the behavior of the model. These values are calculated for all features individually, by utilizing ML model's conditional expected value function [46].

The Shapley value is a concept within cooperative game theory that allocates the overall gains derived from collaboration among participants in a game, based on their respective incremental contributions [45]. It is shown as (1):

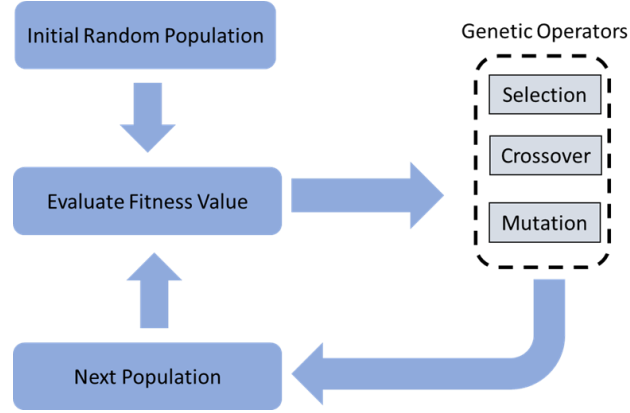


FIGURE 3. Genetic algorithm optimization process.

$$\phi_i(f, x) = \sum_{z' \subseteq x'} \frac{|z'|!(M-|z'|-1)!}{M!} [f_x(z') - f_x(z' \setminus i)] \quad (1)$$

$f$  represents model,  $M$  corresponds to number of  $x'$ , whereas  $x'$  represents a condensed input that corresponds to the initial input via a mapping function  $x = h_x(x')$ . Here  $h_x$  assigns 1 or 0 to the initial inputs, 0 means that input is not considered for the model whereas 1 means otherwise.  $|z'|$  represents the count of non-zero elements within  $z'$ , and  $z' \subseteq x'$  denotes all vectors of  $z'$  where the non-zero elements form a subset of the non-zero elements in  $x'$  [45]. This process calculates a significance value for each feature, and it corresponds to the effect on prediction of the model. The importance of the  $i$ th feature is calculated by comparing the predictions of two models: one trained with all features, and the other trained without the  $i$ th feature. As the impact of excluding a feature is contingent on the presence of other features in the model, the earlier differences are calculated for every potential subset denoted as  $z' \setminus i$ . Thus, the Shapley value can be characterized as a distinctive method for attributing features that computes a weighted average of all potential differences [46].

ML SHAP value is created in a similar manner with the traditional SHAP value, utilizing the conditional expectations to designate mapped inputs [47]. In ML SHAP value calculation, conditional expectation ( $E[f(z)|z_S]$ ) is used instead of  $f_x(z')$  to acquire the SHAP values. Here,  $S$  represents the collection of indices contained within  $z$ . For every sample in the dataset, SHAP values are calculated and therefore contribution of each feature on model prediction can be analyzed [48]. The conventional method for calculating feature importance is to average the absolute

values of the SHAP values that are calculated for all instances. It is shown as (2) ( $N$  represents the number of samples in the dataset):

$$S_0 = \frac{1}{N} \sum_{i=1}^N |\phi_i| \quad (2)$$

### 3. RESULTS

**3.1. Experimental Setup.** The dataset is divided into two parts, one part for training and validation, the other part is for testing. The division ratio is 70% for training and validation, 30% for testing. In GA optimization phase, best model parameters are determined using a tenfold cross-validation (CV) approach. In this approach, the dataset is divided into ten equal parts, with one part reserved for validation while the remaining nine parts serve as the training set. The CV process concludes after each part has been used as the validation set. The CV approach is used to determine the best trained ML model based on available training data. After that the trained model is evaluated on a test set that is not seen by the model before and experimental results are reported afterwards. This train-validate-test approach measures generalization ability of ML models on an unseen test dataset.

Regression performance of ML models are evaluated in terms of Correlation Coefficient (CC), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) metrics. These metrics assess how well the regression models behave when predicting exact Bragg curve peak point location values. To attain a high level of performance, the model should demonstrate low error rates, alongside a high correlation value.

CC is a metric that ranges from -1 to 1. Positive correlation is depicted as +1 whereas negative correlation represented by -1. CC metric is consisting of various variables,  $n$  refers to dataset sample size,  $a_i$  and  $p_i$  variables refer to actual and predicted values, respectively.  $\bar{a}$  and  $\bar{p}$  are calculated mean values of actual and predicted values. The correlation between estimated and actual values is determined using the CC value. It is shown as (3):

$$CC = \frac{S_{PA}}{\sqrt{S_P S_A}} \quad (3)$$

$$S_{PA} = \frac{\sum_i (p_i - \bar{p})(a_i - \bar{a})}{n - 1}$$

$$S_P = \frac{\sum_i (p_i - \bar{p})^2}{n - 1}, \text{ and } S_A = \frac{\sum_i (a_i - \bar{a})^2}{n - 1} .$$

Another metric, MAE quantifies the disparity between two continuous variables. It is given as (4):

$$MAE = \frac{|p_1 - a_1| + \dots + |p_n - a_n|}{n} \quad (4)$$

RMSE, a quadratic metric, serves as a reliable measure of the error magnitude in machine learning models. It effectively quantifies the discrepancy between the model's predicted values and the actual observed values (5):

$$RMSE = \sqrt{\frac{(|p_1 - a_1|)^2 + \dots + |p_n - a_n|^2}{n}} \quad (5)$$

**3.2. Experimental Results.** Various experiments are performed on existing ML models. These models are Decision Tree (DT), Random Forest (RF), Linear Regression (LR), eXtreme Gradient Boosting (XGBoost), Support Vector Regression (SVR) and k-nearest neighbor (kNN). The parameter setup for selected ML models is given in Table 1.

The experimental results for regression models aimed at predicting the peak location values are summarized in Table 2. The RF demonstrated superior performance, achieving correlation coefficient (CC), mean absolute error (MAE), and root mean square error (RMSE) values of 0.712, 0.073, and 0.151, respectively. As can be seen from Table 2, RF model outperforms other regression models across all evaluation metrics. The DT model ranked second in terms of CC and RMSE. XGBoost also ranked second in terms of MAE. In contrast, SVR performed worst in terms of CC, MAE and RMSE values. From these results, we can infer that RF model can accurately map the relationships between biomaterial features and energy levels when predicting Bragg curve peak point locations. Another conclusion that can be drawn from these results is that overall, the models that utilize decision tree architectures performed well when making predictions.

In order to further increase the performance of the best prediction model, metaheuristic GA optimization is applied. In this experiment, GA is used to optimize the hyperparameters of RF. Optimized hyper parameters are number of trees, minimum samples per split, and minimum samples in leaf. Results of this experiment can be seen in Table 3. Different GA population (P) sizes are benchmarked to determine the best optimization approach to tune the hyperparameters of RF. According to Table 3. RF algorithm that is tuned with the population size of 50 achieved highest CC (0.742), lowest MAE (0.069) and RMSE (0.145) values. RF with population size selected as 100 performed worst in terms of CC (0.627), MAE (0.081) and RMSE (0.17). As can be seen from Table 3. GA improved the overall capacity of RF model when predicting the exact values of Bragg curve peak point locations.

TABLE 1. Parameter setup for machine learning models.

Model	Parameters
RF	Number of trees = 100
	Minimum samples per split = 2
	Minimum samples in leaf = 1
	Split criteria = MSE
LR	-
XGBoost	Objective = MSE
	Number of trees=100
	Learning rate = 0.3
	Maximum depth = 6
	Lambda = 1
DT	Split criteria = MSE
	Minimum samples per split = 2
	Minimum samples in leaf = 1
SVR	Kernel = rbf
	C = 1
	Epsilon = 0.1
kNN	k = 3

TABLE 2. Experimental results for regression models aimed at predicting peak value.

Model	CC	MAE	RMSE
<b>RF</b>	<b>0.712</b>	<b>0.073</b>	<b>0.151</b>
LR	0.554	0.103	0.179
XGBoost	0.618	0.078	0.177
DT	0.682	0.081	0.167
SVR	0.517	0.12	0.186
kNN	0.567	0.09	0.178

In order to highlight the positive effect of GA algorithm on RF model, Figure 4 and Figure 5 are given. In Figure 4, it can be seen that in terms CC values, “RF with GA” outperformed “RF without GA”. Same situation is also can be said in terms of error metrics. In Figure 5, “RF with GA” had low MAE and RMSE values compared with “RF without GA”.



TABLE 3. Performance comparison for different population sizes using genetic algorithm.

Model	CC	MAE	RMSE
RF	0.712	0.073	0.151
RF + GA (P. size=10)	0.666	0.077	0.162
RF + GA (P. size=50)	<b>0.742</b>	<b>0.069</b>	<b>0.145</b>
RF + GA (P. size=100)	0.627	0.081	0.17

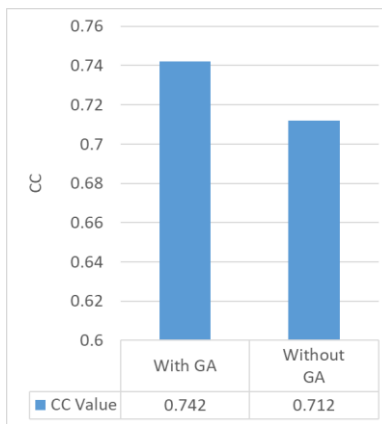


FIGURE 4. Comparison of RF model with applying GA and without GA in terms of CC.

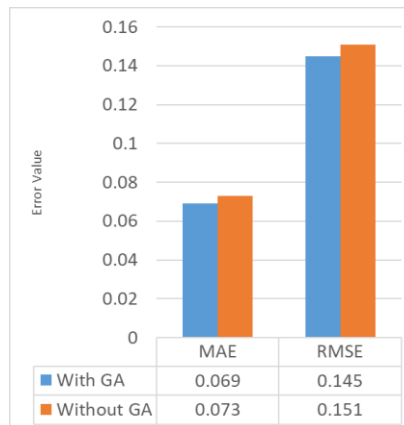


FIGURE 5. Comparison of RF model with applying GA and without GA in terms of MAE and RMSE.

As a final analysis, impact of biomaterial features on Bragg curve peak prediction is evaluated. This evaluation is done by SHAP analysis of “RF + GA (P. size=50)” ML model. In order to demonstrate this analysis, SHAP values of all samples are calculated based on their features.

Summary plots play a crucial role in SHAP analysis, providing not only the hierarchy of input variable importance but also illustrating their correlation with the target variable. One type of summary plot, the beeswarm plot, is illustrated in Figure 6. The beeswarm plot consolidates the SHAP values of all samples, enabling the simultaneous visualization of these values. The summary plot displays input variables and their importance in descending order on the y-axis, while the x-axis represents specific SHAP values. Dot color signifies magnitude (blue for small, red for large), with each dot representing a dataset sample. The horizontal x-axis illustrates the variation in SHAP values for each variable, ranging from blue to red, indicating the shift in input variable magnitude and its impact on prediction. It can be clearly observed that energy, biomaterial mass density, and biomaterial thickness features make a substantial contribution to Bragg curve peak point location prediction. According to Figure 6, lower energy levels result in higher Bragg curve peak point locations, while increased biomaterial mass densities and thickness also lead to higher Bragg curve peak point locations.

It is interesting to note that the model identifies lower energy levels as contributing to higher Bragg curve peak point locations. This seems counterintuitive based on established principles in proton therapy, where lower energy typically leads to shallower peaks [2]. This can be caused by the model might not perfectly capture the complex relationship between energy and Bragg peak location, especially at the lower energy levels. Denser materials require higher energies for protons to reach a specific depth. The model reflects this by predicting a deeper Bragg peak with increasing mass density. This helps tailor the proton beam energy to ensure the peak coincides with the tumor location within the patient's body, composed of tissues with varying densities [3]. Similar to mass density, a thicker material necessitates higher energy protons to achieve a deeper Bragg peak placement. The model's prediction aligns with this principle, allowing for treatment planning that considers the target depth within the patient's specific anatomy. By understanding these feature relationships, the model's predictions can be leveraged to optimize treatment delivery in proton therapy. This translates to more precise targeting of tumors while minimizing radiation exposure to healthy tissues.

#### 4. DISCUSSION AND CONCLUSION

This study aims to utilize ML approaches to accurately predict Bragg curve peak locations in proton beam therapy. The development of a benchmark dataset

encompassing biomaterials that are identical to the human tissue in cancer proton therapies establishes a solid footing for evaluating various ML models. It delves into the effectiveness of different models across varying energy levels and biomaterial characteristics, further employing GA optimization to refine model, thereby bolstering predictive prowess. According to experimental results, the RF model demonstrated superior performance, achieving correlation coefficient (CC), mean absolute error (MAE), and root mean square error (RMSE) values of 0.712, 0.073, and 0.151, respectively. RF model outperforms other regression models across all evaluation metrics. It can be inferred from this result that RF model can map relationships between biomaterial characteristics and energy levels using multiple random tree architectures. Further optimizing RF model parameters using GA is proven to increase the performance of predicting exact value of Bragg curve peak location. RF model with optimized parameters outperformed the model without optimization in terms of 0.742 CC, 0.069 MAE and 0.145 RMSE, respectively.

The study's other main contribution is using SHAP method to dissect the intricate web of feature contributions. This analysis unveils the importance of each feature in predicting Bragg curve peak locations, providing invaluable insights into the factors that profoundly affect treatment outcomes. It can be clearly seen from SHAP analysis that energy, biomaterial mass density, and biomaterial thickness make a substantial contribution to Bragg curve peak point location prediction. Lower energy levels higher Bragg curve peak point locations, while increased biomaterial mass densities and thickness also lead to higher Bragg curve peak point locations. This insight into feature impacts enables personalized adjustments in proton beam therapy, contributing to more tailored and efficacious cancer treatments. The seamless integration of ML model optimization and feature impact analysis not only elevates the precision of proton beam therapy but also represents a substantial leap toward the implementation of individualized and optimized strategies in cancer care.

Through the utilization of these approaches, clinicians can tailor ML models to accommodate diverse patient profiles, optimizing treatment plans to achieve maximum therapeutic impact.

This study has certain limitations, with the primary concern being the computational load. Despite the effectiveness of RF models in various tasks, their computational complexity becomes significant, especially when dealing with a large number of samples and features. The integration of GA optimization can further exacerbate computation time in the quest for the optimal parameter combination. To alleviate this burden, one approach is to employ feature selection and Principal Component Analysis (PCA) methods, which can effectively reduce the dimensionality of features, thereby expediting the training phase.

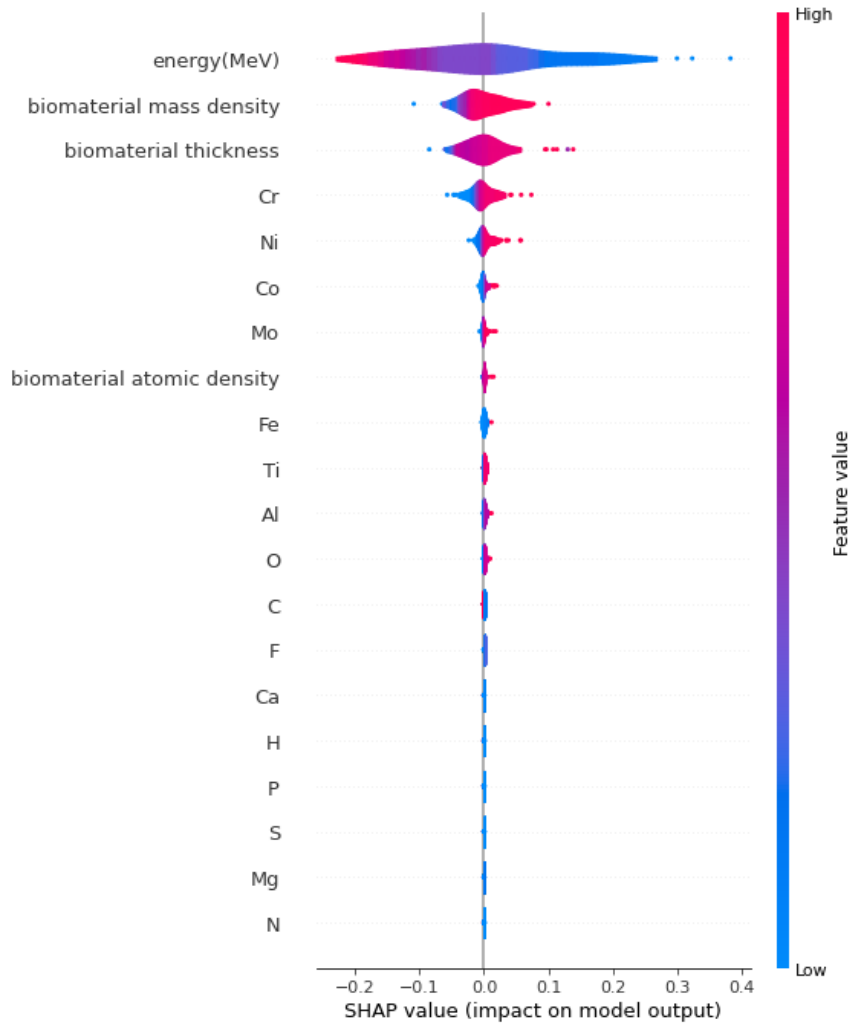


FIGURE 6. Feature importance analysis by SHAP.

An alternative approach involves harnessing Graphics Processor Units (GPUs) within a sophisticated centralized computer system to accelerate the training of the RF algorithm. Additionally, exploring parallel processing techniques can be a viable strategy to address and mitigate this challenge. Another constraint arises from the limited number of samples; given the small sample size, it is impractical to benchmark and study deep learning approaches within the proposed framework using this dataset. Deep learning methods typically require substantial datasets for

effective training. To address these limitations, a potential solution is to conduct additional experiments involving diverse biomaterials and energy levels to augment the sample size. It can increase the performance of proposed framework because the model can train with more varied data. Potential biases within the dataset might not be fully captured by the current analysis. Expanding the dataset with a broader range of biomaterials and energy levels could mitigate bias and improve the model's generalizability to real-world clinical scenarios. While the study demonstrates the effectiveness of the model in predicting Bragg peak location, further work is necessary to translate these findings into clinically applicable tools. Evaluating the model's performance on real-world patient data with complex anatomies would be crucial for establishing its clinical relevance. Also, developing a seamless integration pathway for the model's predictions into existing treatment planning software used by clinicians would enhance its practical usability.

For the future direction of this study, other potential problems, such as dose calculation and treatment planning optimization, which are crucial in particle therapy, will be investigated. The applicability of ML methods to these problems will be explored thoroughly. Various XAI techniques can be explored to understand how these advanced models arrive at their predictions, fostering acknowledgement from medical professionals. Multimodal approaches that consider patient, biomaterial, imaging data can also be integrated for better prediction capability. Especially, multimodal transformers can be adopted to learn joint representations across different data modalities, potentially leading to more comprehensive and informative features for improved prediction capability. Another future task could be increasing the sample size of the curated dataset by conducting more experiments with different biomaterials and energy levels. This approach could open up the possibility of including deep learning methods. Lastly, different hyperparameter optimization algorithms can also be considered to determine which one is better suited for this problem.

**Declaration of Competing Interests** The author declares no known competing interests.

#### REFERENCES

- [1] Ekinci, F., Bölükdemir, M. H., The effect of the second peak formed in biomaterials used in a slab head phantom on the proton Bragg peak, *J. Polytech.*, 23 (1) 2020, 129-136, <http://doi.org/10.2339/politeknik.523001>.
- [2] Ekinci, F., Bostancı, G. E., Dağlı, Ö., Güzel, M. S., Analysis of Bragg curve parameters and lateral straggle for proton and carbon beams, *Commun. Fac. Sci. Univ. Ank. Series*

- A2-A3: *Phys. Sci. and Eng.*, 63 (1) (2021), 32-41, <https://doi.org/10.33769/aupse.864475>.
- [3] Ekinci, F., Bostanci, E., Güzel, M. S., Dağlı, O., Effect of different embolization materials on proton beam stereotactic radiosurgery arteriovenous malformation dose distributions using the Monte Carlo simulation code, *J. Radiat. Res. App. Sci.*, 15 (3) 2022, 191-197, <https://doi.org/10.1016/j.jrras.2022.05.011>.
- [4] Gottschalk, B., Proton Therapy Physics, Taylor & Francis Inc., USA, 2012, <https://doi.org/10.1201/b22053>.
- [5] Ekinci, F., Bostanci, E., Güzel, M. S., Dağlı, Ö., Analysing the effect of a cranium thickness on a Bragg peak range in the proton therapy: a TRIM and GEANT4 based study, *St. Petersburg State Polytech. Univ. J.: Phys. Math.*, 15 (2) (2022) 64-78, <https://doi.org/10.18721/JPM.15207>.
- [6] Carlsson, A. K., Andrea, P. and Brahme, A., Monte Carlo and analytical calculation of computerized treatment plan optimization, *Phys. Med. Biol.*, 42 (1997), 1033-1053, <https://doi.org/10.1088/0031-9155/42/6/004>.
- [7] Hall, E. J., Kellerer, A. M., Rossi, H. H., Lam, Y-M.P., The relative biological effectiveness of 160 MeV protons-II, *Int. Rad. Onc. Biol. Phys.*, 4 (1978), 1009-1013, [https://doi.org/10.1016/0360-3016\(78\)90013-5](https://doi.org/10.1016/0360-3016(78)90013-5).
- [8] Lourenço, A., Wellock, N., Thomas, R., Homer, M., Bouchard, H., Kanai, T., MacDougall, N., Royle, G., Palmans, H., Theoretical and experimental characterization of novel water-equivalent plastics in clinical high-energy carbon-ion beams, *Physics in Medicine and Biology*, 61 (21) (2016), 7623-7638, <https://doi.org/10.1088/0031-9155/61/21/7623>.
- [9] Arib, M., Medjadj, T., Boudouma, Y., Study of the influence of phantom material and size on the calibration of ionization chambers in terms of absorbed dose to water, *J. Appl. Clin. Med. Phys.*, 7 (2006), 55-64, <https://doi.org/10.1120/jacmp.v7i3.2264>.
- [10] Samson, D. O., Jafri, M. Z. M., Shukri, A., Hashim, R., Sulaiman, O., Aziz, M. Z. A., Yusof, M. F. M., Measurement of radiation attenuation parameters of modified defatted soy flour-soy protein isolate-based mangrove wood particleboards to be used for CT phantom production, *Radiat. Environ. Biophys.*, 59 (2020), 483-501, <https://doi.org/10.1007/s00411-020-00844-z>.
- [11] Kanematsu, N., Koba, Y., Ogata, R., Evaluation of plastic materials for range shifting range compensation and solid phantom dosimetry in carbon-ion radiotherapy, *Med. Phys.*, 40 (2013), 041724, <https://doi.org/10.1118/1.4795338>.
- [12] Senirkentli, G. B., Ekinci, F., Bostanci, E., Güzel, M. S., Dağlı, Ö., Karim, A. M., Mishra, A., Therapy for mandibula plate phantom, *Healthcare*, 9 (167) (2021), <https://doi.org/10.3390/healthcare9020167>.
- [13] Ekinci, F., Investigation of tissue equivalence of phantom biomaterials in 4He heavy ion therapy, *Radiat. Eff. Defects Solids*, 178 (3-4) (2023), 500-509,

<https://doi.org/10.1080/10420150.2022.2153251>.

- [14] Ekinci, F., Asuroglu, T., Acici, K., Monte Carlo simulation of TRIM algorithm in ceramic biomaterial in proton therapy, *Materials*, 16 (13) (2023), 4833, <https://doi.org/10.3390/ma16134833>.
- [15] Ekinci, F., Bostanci, E., Güzel, M. S., Dagli, Ö., A Monte Carlo study for soft tissue equivalency of potential polymeric biomaterials used in carbon ion radiation therapy, *Nucl. Technol.*, 209 (8) (2023), 1-11, <https://doi.org/10.1080/00295450.2023.2188144>.
- [16] Borderias-Villarreal, E., et al., Machine learning-based automatic proton therapy planning: Impact of post-processing and dose-mimicking in plan robustness, *Med. Phys.*, 50 (2023), 4480-4490, <https://doi.org/10.1002/mp.16408>.
- [17] Lerendegui-Marco, J., et al., Towards machine learning aided real-time range imaging in proton therapy, *Sci. Rep.*, 12 (2022), 2735, <https://doi.org/10.1038/s41598-022-06126-6>.
- [18] Chang, C. W., Validation of a deep learning-based material estimation model for Monte Carlo dose calculation in proton therapy, *Phys. Med. Biol.*, 67 (21) (2022), 215004, <https://doi.org/10.1088/1361-6560/ac9663>.
- [19] Chen, Y., et al., Understanding machine learning classifier decisions in automated radiotherapy quality assurance, *Phys. Med. Biol.*, 67 (2022), 025001, <https://doi.org/10.1088/1361-6560/ac3e0e>.
- [20] Foster, D. G., Artur, E. D., Average Neutronic Properties of "Prompt" Fission Products, *Los Alamos National Laboratory Report*, LA--9168-MS (1982).
- [21] Ziegler, J. F., SRIM: The stopping and range of ion in matter (2013). Available at: <http://www.srim.org/>. [Accessed November 2023].
- [22] Bhat, P., Malaganve, P., Effect of J48 and LMT algorithms to classify movies in the web-a comparative approach, *Innovations in Computer Science and Engineering. Lecture Notes in Networks and Systems*, Springer Singapore, 2021, [https://doi.org/10.1007/978-981-33-4543-0\\_58](https://doi.org/10.1007/978-981-33-4543-0_58).
- [23] Ilyas, H., et al., Chronic kidney disease diagnosis using decision tree algorithms, *BMC Nephrol.*, 22 (1) 2021, 273, <https://doi.org/10.1186/s12882-021-02474-z>.
- [24] Rahmayanti, N., Pradani, H., Pahlawan, M., Vinarti, R., Comparison of machine learning algorithms to classify fetal health using cardiotocogram data, *Procedia Comp. Sci.*, 197 (2022), 162-171, <https://doi.org/10.1016/j.procs.2021.12.130>.
- [25] Simsekler, M. C. E., Alhashmi, N. H., Azar, E., King, N., Luqman, R., Al Mulla, A., Exploring drivers of patient satisfaction using a random forest algorithm, *BMC Med. Inform. Decis. Mak.*, 21 (1) (2021), 157, <https://doi.org/10.1186/s12911-021-01519-5>.
- [26] Açıcı, K., Erdaş, Ç. B., Aşuroğlu, T., Toprak, M. K., Erdem, H., Oğul, H., A random forest method to detect Parkinson's Disease via gait analysis, *Communications in Computer and Information Science*, Springer, Switzerland, 2017, [https://doi.org/10.1007/978-3-319-65172-9\\_5](https://doi.org/10.1007/978-3-319-65172-9_5).

- [27] Aşuroğlu, T., Oğul, H., A deep learning approach for sepsis monitoring via severity score estimation, *Comput. Methods Programs in Biomed.* 198 (2021), 105816, <https://doi.org/10.1016/j.cmpb.2020.105816>.
- [28] Oyeleye, M., Chen, T., Titarenko, S., Antoniou, G., A predictive analysis of heart rates using machine learning techniques, *Int. J. Environ. Res. Public Health*, 19 (2022), 2417, <https://doi.org/10.3390/ijerph19042417>.
- [29] Huang, L., Song, T., Jiang, T., Linear regression combined KNN algorithm to identify latent defects for imbalance data of ICs, *Microelectron. J.*, 131 (2023), 105641, <https://doi.org/10.1016/j.mejo.2022.105641>.
- [30] Wu, J., et al., Prediction and screening model for products based on fusion regression and XGBoost classification, *Comput. Intell. Neurosc.*, (2022), <https://doi.org/10.1155/2022/4987639>.
- [31] Shin, H., XGBoost regression of the most significant photoplethysmogram features for assessing vascular aging, *IEEE J. Biomed. Health Inform.*, 26 (7) (2022), 3354-3361, <https://doi.org/10.1109/JBHI.2022.3151091>.
- [32] Manoharan, A., Begam, K. M., Aparow, V. R., Sooriamoorthy, D., Artificial neural networks, gradient boosting and support vector machines for electric vehicle battery state estimation: A review, *J. Energy Storage*, 55 (A) (2022), 105384, <https://doi.org/10.1016/j.est.2022.105384>.
- [33] Quan, Q., et al., Research on water temperature prediction based on improved support vector regression, *Neural Comput. & Applic.*, 34 (2022), 8501-8510, <https://doi.org/10.1007/s00521-020-04836-4>.
- [34] Nilashi, M., Abumalloh, R. A., Minaei-Bidgoli, B., Samad, S., Ismail, M. Y., Alhargan, A., Zogaan, W. A., Predicting Parkinson's Disease progression: Evaluation of ensemble methods in machine learning, *J. Health. Eng.*, (2022), <https://doi.org/10.1155/2022/2793361>.
- [35] Sharin, S. N., Radzali, M. K., Sani, M. S. A., A network analysis and support vector regression approaches for visualising and predicting the COVID-19 outbreak in Malaysia, *Health. Anal.*, 2 (2022), 100080, <https://doi.org/10.1016/j.health.2022.100080>.
- [36] Uddin, S., et al., Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction, *Sci. Rep.*, 12 (2022), 6256, <https://doi.org/10.1038/s41598-022-10358-x>.
- [37] Lin, G., Lin, A., Gu, D., Using support vector regression and K-nearest neighbors for short-term traffic flow prediction based on maximal information coefficient, *Infor. Sci.*, 608 (2022), 517-531, <https://doi.org/10.1016/j.ins.2022.06.090>.
- [38] Fayed, H. A., Atiya A. F., Speed up grid-search for parameter selection of support vector machines, *Appl. Soft Comput.*, 80 (2019), 202-210, <https://doi.org/10.1016/j.asoc.2019.03.037>.



- [39] Sun, Y., et al., An improved grid search algorithm to optimize SVR for prediction, *Soft Comput.*, 25 (2021), 5633-5644, <https://doi.org/10.1007/s00500-020-05560-w>.
- [40] Hamdia, K. M., Zhuang, X., Rabczuk, T., An efficient optimization approach for designing machine learning models based on genetic algorithm, *Neural Comput. and Applic.*, 33 (2021), 1923-1933, <https://doi.org/10.1007/s00521-020-05035-x>.
- [41] Da, L., Sun, K., Random forest solar power forecast based on classification optimization, *Energy*, 187 (2019), 115940, <https://doi.org/10.1016/j.energy.2019.115940>.
- [42] Chui, K. T., Gupta, B. B., Vasant, P., A genetic algorithm optimized RNN-LSTM model for remaining useful life prediction of turbofan engine, *Electronics*, 10 (3) (2021), 285, <https://doi.org/10.3390/electronics10030285>.
- [43] Beyaz, S., Açıcı, K., Sümer, E., Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches, *Jt. Dis. Relat. Surg.*, 31 (2) (2020), 175-183, <https://doi.org/10.5606/ehc.2020.72163>.
- [44] Zhou, J., et al., Employing a genetic algorithm and grey wolf optimizer for optimizing RF models to evaluate soil liquefaction potential, *Artif. Intell. Rev.*, 55 (2022), 5673-5705, <https://doi.org/10.1007/s10462-022-10140-5>.
- [45] Lee, Y. G., et al., SHAP value-based feature importance analysis for short-term load forecasting, *J. Electr. Eng. Technol.*, 18 (2023), 579-588, <https://doi.org/10.1007/s42835-022-01161-9>.
- [46] Gramegna, A., Giudici, P., Why to buy insurance? An explainable artificial intelligence approach, *Risks*, 8 (4) (2020), 137, <https://doi.org/10.3390/risks8040137>.
- [47] Kim, Y., Kim, Y., Explainable heat-related mortality with random forest and SHapley Additive exPlanations (SHAP) models, *Sustain. Cities Soc.*, 79 (2022), 103677, <https://doi.org/10.1016/j.scs.2022.103677>.
- [48] Alenezi, R., Ludwig, S. A., Explainability of cybersecurity threats data using SHAP, *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, (2021), 01-10, <https://doi.org/10.1109/SSCI50451.2021.9659888>.

## AUTOMATIC RECOGNITION OF COFFEE BEAN VARIETIES BASED ON PRE-TRAINED CNN ARCHITECTURES

Aynur YONAR<sup>1</sup> and Öznur ÖZALTIN<sup>2</sup>

<sup>1</sup>Department of Statistics, Selcuk University, Konya, TÜRKİYE

<sup>2</sup>Department of Statistics, Atatürk University, Erzurum, TÜRKİYE


**ABSTRACT.** Coffee is an agricultural commodity of fundamental and considerable economic importance on the global market. In this study, the coffee bean varieties were examined from images via artificial intelligence due to their quality and value on the market. This study aims to create an automated system that can efficiently identify coffee beans without requiring a significant amount of time. In this study, five pre-trained Convolutional Neural Network (CNN) architectures were performed to detect four varieties of coffee beans through images. Extracting features from images is a challenging and specialized task. However, CNN possesses the ability to extract features automatically. Therefore, these architectures were employed as both deep feature extractors and classifiers. Primarily, 1600 coffee beans' images were split into 75:25 training and testing sets. Next, 5-fold cross-validation was applied during the training process. This study presented both validation and testing results. Eventually, ShuffleNet achieved the best classification performance with 99.33% and 99.75% accuracy rates in identifying types of coffee beans for the training and testing sets, respectively. As a result, this study has demonstrated that deep learning technologies can automatically recognize the different types of coffee beans.

### 1. INTRODUCTION

Coffee holds a crucial role as an economic crop, significantly influencing global trade and agriculture, and it ranks among the most widely consumed beverages across the globe. Since coffee prices are directly influenced the quality and type of coffee, separating coffee beans is very important for world markets. Assessing the

---

*Keywords.* Artificial intelligence, agricultural economics, classification, coffee beans, convolutional neural networks, cross-validation, pre-trained architectures.

✉ aynursahin@selcuk.edu.tr-Corresponding author;  0000-0003-1681-9398

✉ doanoznur09@gmail.com;  0000-0001-9841-1702.

color of coffee beans plays a vital role in establishing their quality and market value [1, 2].

This evaluation is typically carried out through visual examination or conventional tools, yet these approaches come with challenges like inconsistency, time intensity, and subjectivity. Using computer vision systems instead of traditional methods is an excellent alternative to eliminate these adverse situations. Computer vision systems provide more accurate, impartial, and sensitive classification results [3, 4].

Computer vision is the field of computer science that deals with technologies that enable computers to identify and manipulate objects they see like humans. It covers the subjects of acquiring the image, processing, analyzing, understanding, extracting numerical data from the images, and making decisions.

In recent times, deep learning methods have found extensive application in the field of computer vision and can extract more detailed information than machine learning methods. In deep learning, generally considered a black-box approach, features are automatically determined with input given as images [5]. Then, classification is performed using these features.

The wide range of colors in coffee beans poses a challenge for their classification through visual inspection. Hence, to tackle the classification of coffee beans, various deep learning algorithms, including Convolutional Neural Networks (CNN) architectures, have been suggested in the literature. Some of these are presented below.

Unal *et al.* [6] created a specialized data set containing 1554 images of 3 unique coffee types: Espresso, Kenya, and Starbucks Pike Place coffee beans, and classified them via 4 different CNN-based models: SqueezeNet, Inception V3, VGG16, and VGG19. The findings indicated that SqueezeNet emerged as the most successful model, achieving the highest average classification success rate of 87.3%.

De Oliveira *et al.* [1] proposed a computer vision system based on artificial neural network (ANN) and Bayes classifier to analyze and categorize four green coffee bean types: whitish, cane green, green, and bluish green. The results indicated that the system achieved a 100% accuracy rate in categorizing variations in the color of green coffee beans.

Jumarlis *et al.* [7] provided a website to detect the level of the coffee beans utilizing image input through a web-based program used GLCM (gray-level co-occurrence matrix) and the K-NN (k-Nearest Neighbor) methods, and the system provides 90% accuracy.

Arboleda [8] classified green coffee beans using 22 data mining algorithms consisting of decision tree, discriminant analysis, support vector machine (SVM), K-NN, and ensembles families and obtained the highest classification accuracy with 94.1% by Coarse Tree Algorithm.

Fukai *et al.* [9] developed an automatic coffee bean sorting system for coffee bean producers using deep CNN to detect the type of coffee beans and trained the ANN to implement into Raspberry Pi compute module with the camera module. They compared the results of conventional linear SVM. CNN gives better performance than SVM.

Huang *et al.* [10] designed an automated system for identifying green coffee beans, employing image processing and data augmentation technologies to handle the data, and utilizing CNN to analyze image information. Following the research, a classification accuracy of 94.63% was achieved.

Gope and Fukai [2] classified green coffee bean images through CNN and SVM. Firstly, they created four trained CNN models corresponding to different image sizes, including  $32 \times 32$  pixels,  $64 \times 64$  pixels,  $128 \times 128$  pixels, and  $256 \times 256$  pixels, to compare the classification accuracies of CNN and traditional linear SVM for normal and pea berry coffee beans. The CNN yielded a notable accuracy rate of 96.71%.

Santos *et al.* [11] used SVM, Deep Neural Network (DNN) and Random Forest (RF), to assess defects in coffee beans. They concluded that all classification models performed similarly. In addition to these studies, state-of-the-art studies were also examined, as detailed in Table 1.

The research is focused on efficiently identifying the type of coffee beans through the utilization of deep learning algorithms. In this context, five different CNN-based pre-trained architectures, AlexNet, Inception-v1, MobileNet-v2, ShuffleNet, and SqueezeNet, have been used to classify coffee beans. The advantages and contributions of this study is as follows:

- (i) This study automatically shows recognition of coffee bean types via pre-trained CNN.
- (ii) Expert opinion is not needed to extract features.
- (iii) To find the optimal architecture, AlexNet, Inception-v1, MobileNet-v2, ShuffleNet, and SqueezeNet are compared by using distinct performance metrics on training and testing sets.
- (iv) ShuffleNet has the highest performance on training and testing set to determine coffee beans.
- (v) This study may shed light on the determination of the quality and diseases of agricultural products.

The pipeline of this study is shown in Figure 1. The performance of the techniques has been compared using AUC, accuracy, sensitivity, specificity, precision, F1-score, G-Mean metrics, and ROC curves. Finally, the model having the highest success is determined.

TABLE 1. State-of-the-art studies.

Study	Coffee Classes	Methods	Metrics	
Tsai <i>et al.</i> [12]	2 classes	Mass spectrometry (MS) analysis + ANN	Accuracy	0.9958
			Sensitivity	0.9875
			Specificity	1.0000
Arboleda [13]	2 classes	Feature extraction + K-NN	Accuracy	0.9700
Raveena and Surendran [14]	6 classes	ResNet50	Accuracy	0.9897
			Sensitivity	0.9844
			F1-score	0.9864
		VGG16	Accuracy	0.9638
			Sensitivity	0.9523
			F1-score	0.9563
Kim <i>et al.</i> (2024) [15]	2 classes	CNN Based Model	Accuracy	0.9927
Chang and Liu (2024) [16]	8 classes	CNN Based Model	Accuracy	0.9600
			Kappa	0.9500
		VGG-16	Accuracy	0.8100
			Kappa	0.7900
		ResNet-18	Accuracy	0.8900
			Kappa	0.8490
		AlexNet	Accuracy	0.8900
			Kappa	0.8000
		GoogleNet	Accuracy	0.9200
			Kappa	0.9000

The remainder of this paper is organized as follows: In the "Material and Methods" section, details about the coffee beans dataset, CNN and pre-trained architectures, cross-validation, confusion matrix, and performance measures are provided. The "Experimental Results" section presents the outcomes, while the "Conclusion" section offers concluding remarks.

## 2. MATERIALS AND METHODS

This section provides concise information about the coffee beans dataset, CNN, pre-trained architectures, cross-validation, and performance measures utilized in the study.

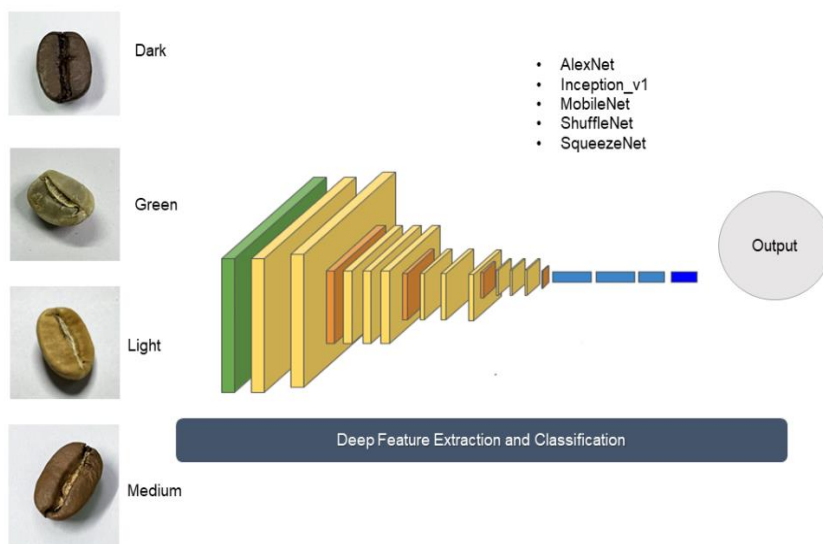


FIGURE 1. Pipeline of this study.

**2.1. Coffee Beans Dataset.** The coffee beans dataset used in this study were from study of [17]. In addition, the dataset is publicly available on the Kaggle platform. There are four different coffee classifications: dark, green, light, and medium. There are 1600 images in total, 400 in each class. Each example bean's image is 224x224x3 pixels in size, and images of four different coffee beans were used to recognize the coffee type via deep learning. This study's dataset is split into 75% training and 25% testing sets.

**2.2. Convolutional neural network (CNN).** The convolutional neural network (CNN) stands out as a prominent deep learning method characterized by its intricate structure composed of multiple layers. CNN are commonly applied to address image processing challenges due to their capability to conduct feature extraction, learning, and classification based on these extracted features. Moreover, CNN overcomes the computational complexity problem that other classification algorithms have in real-time data and provides very good classification results in studies involving both large and small datasets.

CNN processes an image in different layers and then extracts all its features. The most used layers are given below [6, 18, 19]:

**Convolution Layer:** This layer serves as the fundamental building block of CNNs, extracting features by systematically applying different filters to the image. To effectively determine the increasing number of features, it is essential to augment both the number of steps and filters in conventional layers at a proportional rate. However, as the number of features increases, learning becomes more difficult for the network, so this number must be determined optimally.

**Pooling Layer:** This layer simplifies big data from the convolution layer by preserving their existing properties to reduce programming complexity and improve learning.

**Activation Layer:** This layer, also known as the non-linear layer, activates the system with non-linear functions and prevents values from falling outside the valid data range.

**Fully Connected Layer:** This layer is the crucial artificial neural network layer within CNNs, playing a pivotal role in the learning processes and feature extraction.

**SoftMax Layer:** The distribution of classes becomes apparent, and an output is generated through the labeling process within this layer.

In the context of this study, five distinct pre-trained CNN architectures—AlexNet, Inception\_v1, MobileNet\_v2, ShuffleNet, and SqueezeNet were employed for the classification of coffee bean varieties.

**AlexNet** was proposed by [20] for image classification. The network won the ImageNet Large-Scale Visual Recognition Competition (ILSVRC) 2012 with more than 26% accuracy over contemporary models. It comprises eight trainable layers, including five convolutional layers and three fully connected layers. The last layer of the fully connected layer is associated with a SoftMax classifier configured for N classes, where N denotes the number of classes. The network employs multiple convolutional kernels for the extraction of features from the image. Additionally, it incorporates dropout for regularization and a rectified linear unit (ReLU) activation function to expedite training convergence [21].

**Inception\_v1** model, which was the winner of ILSVRC in 2015, has a significant success in the development of CNN classifiers. Convolution occurs with three sizes of filters ( $1 \times 1$ ,  $3 \times 3$ ,  $5 \times 5$ ) at the same level with maximum pooling. The outputs from this more comprehensive layer are combined and fed as input to the next layer.

**MobileNet\_v2** was proposed by Sandler et al. in 2018. MobileNetv2 comprises one convolutional layer, succeeded by 19 residual bottleneck modules, and subsequently, two convolutional layers. The bottleneck module incorporates a shortcut connection exclusively when the stride is set to 1. Shortcut is not used for higher pitch due to size difference. They also used ReLU as a nonlinear function instead of simple ReLU to limit the calculations.

**ShuffleNet**, introduced by [22], is primarily composed of a standard convolution and a series of ShuffleNet units organized into three stages. The ShuffleNet unit bears resemblance to the ResNet block, employing depth convolution on 3x3 layers and substituting the 1x1 layer with point group convolution. The depth convolution layer is preceded by a channel blending process.

**SqueezeNet** stands out as a more compact and innovative CNN architecture, characterized by fewer parameters compared to other CNN models. SqueezeNet consists of fifteen layers, including two convolution layers, three max-pooling layers, eight fire layers, a global average pooling layer, and a SoftMax layer with an output layer. Fire layers create compression and expansion between convolution layers. SqueezeNet is an excellent candidate to improve the hardware efficiency of neural network architectures. Details of pre-trained architectures used in this study are given in Table 2.

**2.3. Cross-Validation.** During the separation of the data set into the training and test sets, irregular distribution of the data set may negatively affect the model's performance. This problem can be solved with the k-fold cross-validation method. The dataset is partitioned into segments represented as  $k$  folds in cross-validation. Subsequently,  $k-1$  folds are trained in the framework and tested on the remaining folds at each step. The critical point here is to use the previously untested part as the test set in each step [24]. We used 5-fold cross-validation for train data in this study.

TABLE 2. Details of pre-trained architectures used in this study [23].

Pre-trained architectures	Features		
	Parameters (millions)	Input Image Size	Depth
AlexNet	61	227x227	8
Inception_v1	7	224x224	22
MobileNet_v2	3.5	224x224	53
ShuffleNet	1.4	224x224	50
SqueezeNet	1.24	227x227	18



2.4. **Confusion matrix and Performance measures.** The confusion matrix illustrates the current state of the dataset, presenting the count of both accurate and inaccurate predictions made by the classification model in a tabular format. For the evaluation of classification model performance in this study, a four-class confusion matrix was utilized. Moreover, six performance metrics derived from the confusion matrix were utilized to analyze the results of the experimental study, as detailed in Table 3. Additionally, the distinctiveness of the results was assessed using the values of the receiver operating characteristic (ROC) curves and AUC (area under the curve). In the ROC curve, the false positive rate is represented on the x-axis, while the true positive rate is depicted on the y-axis. An AUC value approaching 1 signifies high classification performance for the method [25].

TABLE 3. Performance metrics formulas. *TP* : True Positive, *TN* : True Negative, *FP* : False Positive, *FN* : False Negative.

Metric	Formula
Accuracy	$ACC = \frac{TP + TN}{TP + TN + FP + FN}$
Sensitivity	$TPR = \frac{TP}{TP + FN}$
Specificity	$TNR = \frac{TN}{TN + FP}$
Precision	$PPV = \frac{TP}{TP + FP}$
F1-Score	$F1\ score = \frac{2}{\frac{1}{TPR} + \frac{1}{PPV}}$
G-Mean	$\sqrt{TPR \times TNR}$

### 3. EXPERIMENTAL RESULTS

Section-Times New Roman 11, Figure -illustrated in Figure 1. This study focuses on detecting coffee bean types via artificial intelligence. The acquired images of coffee beans were categorized employing five distinct pre-trained CNN models: AlexNet, Interception v1, MobileNet\_v2, ShuffleNet, and SqueezeNet. The models were trained in the MATLAB environment, utilizing an Intel Core i7-7500U CPU, NVIDIA GeForce GTX 950M, 16 GB RAM, and a 64-bit Operating System.

The dataset was divided into 75% train and 25% test sets. 5-fold cross-validation was applied to the training set to obtain confident results.

The confusion matrix of each CNN model is presented in Figure 2. There are correct and incorrect classification numbers for each algorithm. For example, the AlexNet model correctly classified 92, 46, 97, and 95 images of Dark, Green, Light, and Medium coffee beans, respectively. The category with the highest misclassification of coffee bean images is the green coffee bean class.

The performance of the models was evaluated using metrics including accuracy, sensitivity, specificity, precision, recall, and F-1 score, along with G-Mean. Table 4 and 5 show 5-fold cross-validation and test results for each algorithm, respectively. All performance matrices achieved the highest value with ShuffleNet for validation. However, when the architectures were tested, the best optimal performance was obtained with MobileNetv2 and ShuffleNet to detect coffee bean varieties. As a result, ShuffleNet was found to be the most efficient net in terms of validation and testing data results. Because the difference between the validation and test results is as small as possible, the algorithm does not learn excessively. Therefore, it was the best one. In Table 4 and Table 5, bold values display the highest performance metrics.

To have information about the distinctiveness of the models, ROC curves are drawn and given in Figure 3. It is observed that the ShuffleNet model exhibits the highest level of distinctiveness.

TABLE 4. Average performance metrics of all models for validation data.

CNN Algorithm	AUC	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-Mean
AlexNet	0.9717	0.8025	0.8025	0.9342	0.8270	0.8043	0.8658
Inception_v1	0.9992	0.9892	0.9892	0.9964	0.9894	0.9892	0.9928
MobileNet_v2	0.9995	0.9842	0.9842	0.9947	0.9847	0.9842	0.9894
ShuffleNet	<b>0.9999</b>	<b>0.9933</b>	<b>0.9933</b>	<b>0.9978</b>	<b>0.9933</b>	<b>0.9933</b>	<b>0.9956</b>
SqueezeNet	0.9972	0.9592	0.9592	0.9864	0.9593	0.9591	0.9727

TABLE 5. Average performance metrics of all models for test data.

CNN Algorithm	AUC	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-Mean
AlexNet	0.9918	0.8250	0.8250	0.9417	0.8796	0.8179	0.8814
Inception_v1	<b>1.0000</b>	0.9950	0.9950	0.9983	0.9951	0.9950	0.9967
MobileNet_v2	<b>1.0000</b>	<b>0.9975</b>	<b>0.9975</b>	<b>0.9992</b>	<b>0.9975</b>	<b>0.9975</b>	<b>0.9983</b>
ShuffleNet	<b>1.0000</b>	<b>0.9975</b>	<b>0.9975</b>	<b>0.9992</b>	<b>0.9975</b>	<b>0.9975</b>	<b>0.9983</b>
SqueezeNet	0.9998	0.9675	0.9675	0.9892	0.9706	0.9679	0.9783

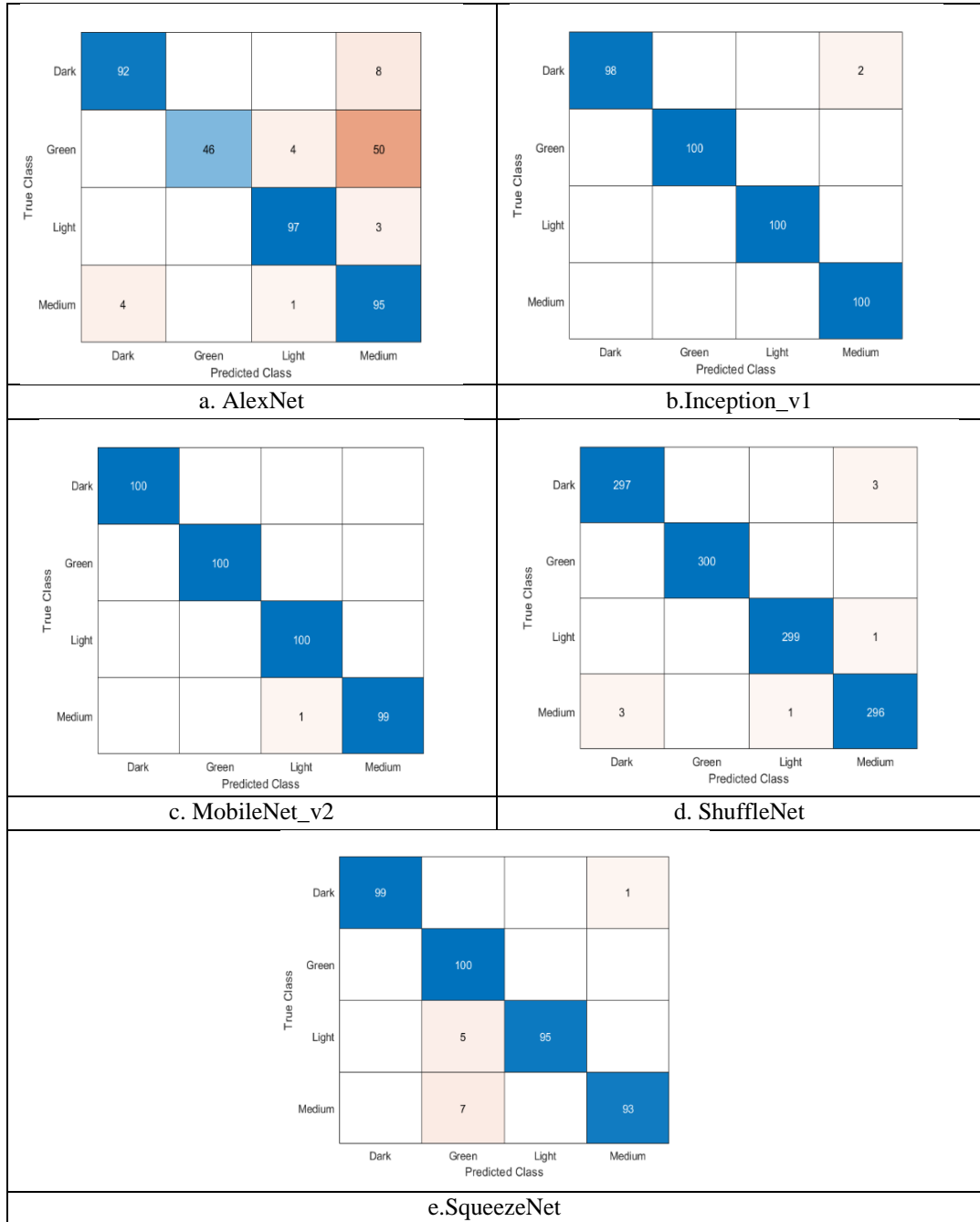


FIGURE 2. Confusion matrixes of CNN algorithms.

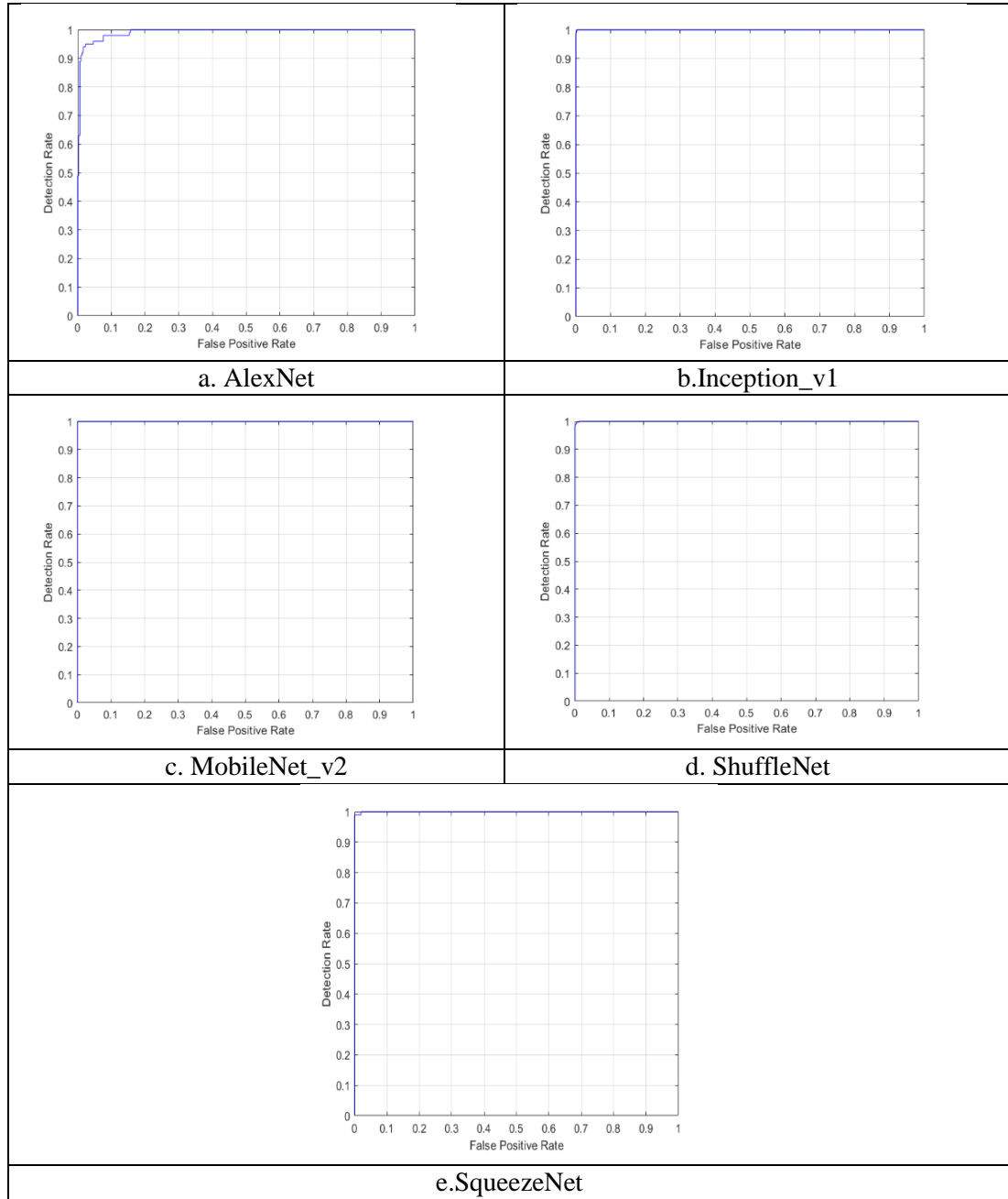


FIGURE 3. ROC Curves of CNN algorithms.

## 5. CONCLUSIONS

In this study, five distinct CNN architectures—AlexNet, Inception\_v1, MobileNet\_v2, ShuffleNet, and SqueezeNet—were employed for the classification of coffee beans. The coffee bean dataset was initially split into a training set comprising 75% of the data and a testing set consisting of the remaining 25%. Additionally, a training set and validation set were created to prevent overfitting by applying 5-fold cross-validation to the training set. For the validation set, the best architecture, according to all performance metrics, was determined to be ShuffleNet. Among the testing set utilized for assessing model performance, MobileNet\_v2 and ShuffleNet demonstrated the highest success in classifying the test set. However, it is expected that the difference between the validation or training result and the test results will be minimal. In this study, Shufflenet was identified with the smallest difference as the top choice. Thus, the utilization of ShuffleNet architecture in detecting coffee beans can streamline quality control processes and minimize decision-making errors.

This study has some constraints:

- (i) Types of coffee bean is limited. In addition to the identification of species, diseases could also be identified.
- (ii) If the same images had features extracted by experts, comparisons could be made with other artificial intelligence algorithms.

In future studies, different existing or newly proposed architectures and other coffee bean or agricultural image datasets can be used for comparison purposes.

**Author Contribution Statements** The authors contributed equally to this work.

**Declaration of Competing Interests** The authors declare that there is no conflict of interest regarding the publication of this paper.

## REFERENCES

- [1] De Oliveira, E. M., Leme, D. S., Barbosa, B. H. G., Rodarte, M. P., Pereira, R. G. F. A., A computer vision system for coffee beans classification based on computational intelligence techniques, *J. Food Eng.*, 171 (2016), 22-27.
- [2] Gope, H. L., Fukai, H., Normal and pea berry coffee beans classification from green coffee bean images using convolutional neural networks and support vector machine, *Int. J. Comput. Inf. Eng.*, 14 (6) (2020), 189-196.
- [3] Adiwijaya, N. O., Romadhon, H. I., Putra, J. A., Kuswanto, D. P., The quality of coffee bean classification system based on color by using k-nearest neighbor method, *J. Phys.: Conference Series*, 2157 (2022).

- [4] Vogt, M. A. B., Developing stronger association between market value of coffee and functional biodiversity, *J. Environ. Manage.*, 269 (2020).
- [5] Buhrmester, V., Münch, D., Arens, M., Analysis of explainers of black box deep neural networks for computer vision: *A survey, Mach. Learn. Knowl. Extr.*, 3 (4) (2021), 966-989.
- [6] Unal, Y., Taspinar, Y. S., Cinar, I., Kursun, R., Koklu, M., Application of pre-trained deep convolutional neural networks for coffee beans species detection, *J. Food Anal. Method*, 15 (12) (2022), 3232-3243.
- [7] Jumarlis, M., Mirfan, M., Manga, A. R., Classification of coffee bean defects using gray-level co-occurrence matrix and k-nearest neighbor, *ILKOM J. Ilmiah*, 14 (1) (2022), 1-9.
- [8] Arboleda, E. R., Comparing performances of data mining algorithms for classification of green coffee beans, *J Int. J. Eng. Adv. Technol*, 8 (5) (2019), 1563-1567.
- [9] Fukai, H., Furukawa, J., Katsuragawa, H., Pinto, C., Afonso, C., Classification of green coffee beans by convolutional neural network and its implementation on raspberry Pi and Camera Module, *Timor. Acad. J. Sci.*, 1 (2018), 1-10.
- [10] Huang, N. F., Chou, D. L., Lee, C. A., Wu, F. P., Chuang, A. C., Chen, Y. H., Tsai, Y. C., Smart agriculture: real-time classification of green coffee beans by using a convolutional neural network, *JJET Smart Cities*, 2 (4) (2020), 167-172.
- [11] Santos, F. F. L. d., Rosas, J. T. F., Martins, R. N., Araújo, G. d. M., Viana, L. d. A., Gonçalves, J. d. P., Quality assessment of coffee beans through computer vision and machine learning algorithms, *Coff. Sci.*, (2020).
- [12] Tsai, J.-J., Chang, C.-C., Huang, D.-Y., Lin, T.-S., Chen, Y.-C., Analysis and classification of coffee beans using single coffee bean mass spectrometry with machine learning strategy, *Food Chem.*, (2023), 426, <https://doi.org/10.1016/j.foodchem.2023.136610>.
- [13] Arboleda, E., Classification of immature and mature coffee beans using texture Features and medium K nearest neighbor, *J. Artif. Intell. Technol*, 3 (3) (2022), 114-118, <https://doi.org/10.37965/jait.2023.0203>.
- [14] Raveena, S., Surendran, R., ResNet50-based classification of coffee cherry maturity using deep-CNN, *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, (2023), 1275-1281, <https://doi.org/10.1109/ICSSIT55814.2023.10061006>.
- [15] Kim, Y., Lee, J., Kim, S., Study of active food processing technology using computer vision and AI in coffee roasting, *Food Sci. Biotechnol.*, (2024), 1-8, <https://doi.org/10.1007/s10068-023-01507-7>.
- [16] Chang, S.-J., Liu, K.-H., Multiscale defect extraction neural network for green coffee bean defects detection, *IEEE Access*, 12 (2024), 15856-15866, <https://doi.org/10.1109/ACCESS.2024.3356596>.

- [17] Ontoum, S., Khemanantakul, T., Sroison, P., Triyason, T., Watanapa, B., Coffee roast intelligence, arXiv: 2206.01841, (2022).
- [18] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M. S., Deep learning for visual understanding: A review, *Neurocomputing*, 187 (2016), 27-48.
- [19] Ozaltin, O., Yeniay, O., A novel proposed CNN–SVM architecture for ECG scalograms classification, *Soft Comput.*, 27 (8) (2023), 4639-4658.
- [20] Krizhevsky, A., Sutskever, I., Hinton, G. E., Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.*, 25 (2012).
- [21] Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M., Lee, B., A survey of modern deep learning-based object detection models, *Dig. Signal Process.*, 126, (2022).
- [22] Zhang, X., Zhou, X., Lin, M., Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2018).
- [23] Mathworks, 2023, <https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html>.
- [24] Yadav, S., Shukla, S., Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification, *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, (2016).
- [25] Gorunescu, F., Gorunescu, F., Classification performance evaluation, *Data Mining: Concepts, Models Techniques*, (2011), 319-330.

## HEMOGLOBIN VALUE PREDICTION WITH BAYESIAN OPTIMIZATION ASSISTED MACHINE LEARNING MODELS

Koray AÇICI<sup>1</sup>


<sup>1</sup>Department of Artificial Intelligence and Data Engineering, Ankara University,  
Ankara, TÜRKİYE

**ABSTRACT.** This study presents a framework for predicting hemoglobin (Hb) levels utilizing Bayesian optimization-assisted machine learning models, incorporating both time-domain and frequency-domain features derived from photoplethysmography (PPG) signals. Hemoglobin, a crucial protein for oxygen and carbon dioxide transport in the blood, has levels that indicate various health conditions, including anemia and diseases affecting red blood cell production. Traditional methods for measuring Hb levels are invasive, posing potential risks and discomfort. To address this, a dataset comprising PPG signals, along with demographic data (gender and age), was analyzed to predict Hb levels accurately. Our models employ support vector regression (SVR), artificial neural networks (ANNs), classification and regression trees (CART), and ensembles of trees (EoT) optimized through Bayesian optimization algorithm. The results demonstrated that incorporating age and gender as features significantly improved model performance, highlighting their importance in Hb level prediction. Among the tested models, ANN provided the best results, involving normalized raw signals, feature selection, and reduction methods. The model achieved a mean squared error (MSE) of 1.508, root mean squared error (RMSE) of 1.228, and R-squared ( $R^2$ ) of 0.226. This study's findings contribute to the growing body of research on non-invasive hemoglobin prediction, offering a potential tool for healthcare professionals and patients for convenient and risk-free Hb level monitoring.

### 1. INTRODUCTION

The iron-rich protein hemoglobin (Hb), found inside red blood cells, is crucial for transporting oxygen and carbon dioxide throughout the body. Fundamentally, if Hb level is low, tissues cannot obtain necessary oxygen [1]. Hemoglobin levels play a

*Keywords.* Hemoglobin, machine learning, regression.

✉ kacici@ankara.edu.tr -Corresponding author;  0000-0002-3821-6419.



vital role in overall health, and abnormal levels can indicate underlying diseases. Especially, low Hb levels are a sign of anemia, a condition where the body doesn't have enough healthy red blood cells. Since iron is fundamental to produce hemoglobin, insufficient iron may lead to iron-deficiency anemia, which is the most prevalent form [2]. The shape of red blood cells is determined genetically. If an individual has a sickle cell disease, sickle-shaped red blood cells can block blood vessels, causing pain and tissue damage [3, 4]. Another genetic disorder, known as Thalassemia, affects the production of hemoglobin, resulting in lower levels than normal in the bodies of individuals [5]. It is known that some types of leukaemia can also affect the production of red blood cells, leading to anemia [6]. Healthy kidneys secrete a hormone known as erythropoietin (EPO), which aids in the stimulation of red blood cell creation. When the kidneys are damaged or in chronic kidney disease, they may not produce enough EPO and this situation leads to anemia [7].

According to the WHO guidelines anemia classification is based on hemoglobin levels [8]. Acceptable hemoglobin levels show difference for men and women. In mild anemia hemoglobin value is between 11 g/dL and 12.9 g/dL for men and between 11 g/dL and 11.9 g/dL for women. In moderate anemia, hemoglobin value is between 8 g/dL and 10.9 g/dL for both men and women. In severe anemia hemoglobin value is lower than 8 g/dL for both men and women. Therefore, the hemoglobin levels play an important role in people's lives. Especially for patients with hemoglobin related diseases, it is vital to measure hemoglobin values.

Traditionally, hemoglobin levels are measured through blood tests, which can be inconvenient. Invasive methods for measuring hemoglobin concentration (Hb) are generally safe, but there are some potential risks involved. Some low risks include the pain caused by the needle to draw blood, slight dizziness, light-headedness, and bruising at the puncture site. However, there are some serious risks associated with invasive Hb measurement, the probability of them occurring is very low. Infection, excessive bleeding, and fainting can be given as examples of potential serious risks. Machine learning offers inspiring possibilities for non-invasive hemoglobin prediction through various techniques. Leveraging the ubiquity of smartphones, researchers have developed hemoglobin prediction tools using smartphone cameras and built-in light sources. Techniques involve analyzing fingertip images or videos, focusing on color variations related to blood oxygenation [9]. Photoplethysmography (PPG) is a widely explored technique that uses light to measure blood volume changes in tissues. Machine learning algorithms can analyze features extracted from the PPG signal, such as pulse rate and amplitude, to predict hemoglobin levels [10, 11].

In this study, a machine learning framework was proposed including PPG signals. Our contribution to literature is two-fold. To the best of our knowledge, the dataset published by Abuzairi et al. was utilized in this study for the first time [12]. Second,

a Bayesian optimized method was applied in the training phases of machine learning algorithms.

The organization of the article is as follows: In Section 2, the related works are summarized. In Section 3, materials and methods including the dataset, machine learning, feature selection, feature reduction, and optimization algorithms are described. In Section 4, the experimental setup, the evaluation metrics, and the empirical results are presented. Section 5 concludes the article.

## 2. RELATED WORKS

There's been a growing focus on using machine learning for non-invasive hemoglobin prediction. The study by Dimauro et al. [13] proposed a non-invasive method for estimating hemoglobin (Hb) concentration based on digital images of the conjunctiva. This innovative approach aimed to assess anemia without requiring a blood sample, making it more convenient for patients and healthcare providers. Their prototype extracts essential information from colour values in acquired images of the conjunctiva. Participants were mainly recruited from Hematology Departments and a transfusion center in Italy. Each subject allowed one blood sample for laboratory Hb measurement, and simultaneously, images of their conjunctiva were acquired using the proposed device. Tests on a mix of 113 anemic and healthy individuals demonstrated a strong correlation between the device's estimated Hb value and the actual Hb value. A k-nearest neighbor (kNN) classification algorithm was employed to assess the anemic condition based on features extracted from the conjunctiva images. The study utilized the CIE  $L^*a^*b^*$  color space for image analysis, focusing on extracting mean values of the  $a^*$ ,  $b^*$  components, and the G value from the RGB components of the conjunctiva images. The methodology included filtering input data based on lightness (L) and RGB components to exclude pixels that were too dark or too bright, ensuring that only pixels allowing correct pallor evaluation were considered. Pearson Correlation Index between conjunctival  $a^*$  mean values and measured Hb was found to be 0.726 for the full dataset indicating a strong correlation. The authors concluded that their proposed method and device could serve as an effective tool for non-invasive anemia screening and monitoring, with the potential for use both in medical settings and by patients at home.

Another study presented a method for the non-invasive diagnosis of anemia through Hb detection using a spectrophotometric system and a BP-ANN model [14]. In their study, the dataset consists of fingertip spectra from 109 volunteers, with 4 samples identified as outliers and removed, leaving 105 samples for the analysis. Samples were divided into calibration (53 samples), correction (26 samples), and prediction (26 samples) sets. A spectrophotometric system was developed, incorporating a broadband light source, grating spectrograph, and silicon photodiode

array for measuring the fingertip spectra. Principal Component Analysis (PCA) was employed to reduce the dimensionality of the collected spectra and eliminate redundant data. The principal components were then used as inputs to the BP-ANN model, with the optimal network structure having 9 input nodes (corresponding to the principal components), 11 hidden nodes, and 1 output node. The BP-ANN model was trained and validated using the calibration and correction sample sets, respectively, and tested with the prediction sample set. The correlation coefficient (CC) of the BP-ANN model established by this method was 0.94, indicating a strong correlation between the predicted and actual Hb levels. The study successfully demonstrated the feasibility of non-invasively predicting hemoglobin levels using a combination of PCA and BP-ANN, with satisfactory accuracy and robustness. However, the article did not explicitly provide metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), or the coefficient of determination ( $R^2$ ) for the BP-ANN model's predictions, which are commonly used to evaluate the performance of regression models.

A study, demonstrating the feasibility of measuring hemoglobin levels noninvasively using a standard smartphone's built-in RGB camera and white LED flash, was presented by Wang et al. [15]. The study involved 32 participants, providing a dataset for evaluating the proposed hemoglobin measurement system. Hemoglobin levels were compared against measurements, taken by a device known for optical hemoglobin measurement, to validate the smartphone-based approach. The proposed system extracted features from the PPG signals, focusing on the ratio of peak to trough intensities across different wavelengths (color channels), to assess blood absorption characteristics indicative of hemoglobin levels. A linear regression model that correlates the features extracted from the PPG signals to hemoglobin levels was employed in the study. Although the document did not specifically mention feature reduction techniques, it highlighted the importance of adjusting color channel gain to balance signal contributions from each channel, effectively optimizing the feature set for regression analysis. A Pearson correlation of 0.62 with the reference device was reported, indicating a moderate positive correlation between the smartphone-based measurements and the reference hemoglobin levels. Additionally, an RMSE value of 1.27 g/dL demonstrated the typical deviation of the smartphone-based hemoglobin estimates from the reference measurements.

A comprehensive study on non-invasively predicting hemoglobin levels using PPG signals and various machine learning algorithms was contributed by Kavsaoğlu et al. [16]. The dataset included data from 33 individuals. PPG signals were collected for each participant over 10 periods. Additionally, gender, height, weight, and age were added as features, which increased the total number of features to 44. Hemocue Hb-201TM device was utilized simultaneously with PPG signal collection as a reference for Hb levels. In the study Classification and Regression Trees (CART),

Least Squares Regression (LSR), Generalized Linear Regression (GLR), Multivariate Linear Regression (MVLRL), Partial Least Squares Regression (PLSR), Generalized Regression Neural Network (GRNN), Multilayer Perceptrons (MLP), and Support Vector Regression (SVR) machine learning algorithms were utilized to predict Hb levels. 40 characteristic features were derived from the PPG signal, including time-domain features from the signal and its first and second derivatives. RELIEF based feature selection (RFS) and Correlation-based feature selection (CFS) were utilized to reduce feature dimensions to 10 and 11 features, respectively. As performance metrics, MAE, MSE,  $R^2$ , RMSE, Mean Absolute Percentage Error (MAPE), and Index of Agreement (IA) were taken into consideration to calculate the effectiveness of the algorithms. RFS-assisted SVR provided promising results with the lowest MSE of 0.0027. The study demonstrated that machine learning techniques could effectively predict hemoglobin levels non-invasively using PPG signals and selected characteristic features, offering a viable method for continuous, pain-free monitoring of hemoglobin levels.

Another study that exploited Artificial Neural Network (ANN) architecture to focus on developing a non-invasive method for estimating blood hemoglobin levels was presented by Hasan et al. [17]. Their study involved 75 adults, with hemoglobin levels ranging from 7.6 to 13.5 g/dL. The data collection was performed by using 10-second fingertip videos recorded with a smartphone, resulting in 300 frames per video. The participants' ages ranged from 20 to 56 years. For feature extraction, RGB pixel intensities were obtained from 100 area blocks in each frame. Then, ANN was utilized to build a prediction model for hemoglobin values. A correlation rank order of 0.93 between the predicted hemoglobin values by the model and the gold standard was noted, signifying a high level of predictive accuracy. Additionally, the dataset was divided into 2 categories for classification purposes. Finally, the proposed method demonstrated 94% sensitivity and 96% specificity performance.

El-Kenawy et al., presented a study on using machine learning techniques for estimating Hemoglobin levels and classifying Anemia based on hematological parameters [18]. Their dataset consisted of 9004 records, which were split into training (75%) and testing (25%) data. The training dataset included 6753 records, while the testing dataset had 2251 records. Z-score Normalization was applied for standardizing the data. Some parameters like gender and age were omitted due to incomplete data. ANN, LR, and Random Forest (RF) regressors were employed to estimate Hb levels. The RF model outperformed other regression models in estimating Hemoglobin levels with the lowest RMSE (0.0123) and MAE (0.0435). For anemia classification, several classifiers were tested. A hybrid classifier combining Decision Tree (DT), Naive Bayes (NB), and ANN, optimized through weighted average probabilities, obtained the best performance with an RMSE value of 0.0838 and a MAE value of 0.0159. The study demonstrated that machine learning

techniques, particularly ensemble methods like RF for regression and a hybrid model for classification, can effectively estimate Hemoglobin levels and classify anemia types.

A novel approach for estimating Hb levels non-invasively by using PPG signals captured at four different wavelengths was presented by Chen et al. [19]. Their dataset consisted of 58 volunteers, aged between 21-27, with an approximately equal male-to-female ratio. The signals were collected at a 200 Hz sampling rate for 1 minute. For the preprocessing stage, a second-order Butterworth bandpass filter was implemented to process the raw PPG signal, removing high-frequency noise and motion artifacts. 160 morphological and time-domain feature parameters from the PPG signal across four channels were extracted. To identify the most relevant features, reliefF, Chi-square Score, and Information Gain methods were employed. Three machine learning algorithms, Logistic Regression (LR), SVR, and eXtreme Gradient Boosting (XGBoost), were utilized to obtain models. The XGBoost model, utilizing the top 30 features selected via the Chi-square method, achieved the best performance with a  $R^2$  value of 0.997, a RMSE value of 0.762, and a MAE value of 0.325. The utilization of XGBoost, in combination with carefully selected PPG signal features, represented a novel contribution to the field of non-invasive hemoglobin measurement, showcasing the potential for clinical application.

Another approach to non-invasively predict Hb concentrations by using PPG signals was contributed by Peng et al. [10]. The research included 249 volunteers, with 199 samples allocated to a training set and 50 samples to a test set. An eight-wavelength PPG signal acquisition system, alongside a reference value of Hb concentration from an automatic blood cell analyzer were utilized for data collection. 56 feature values were extracted from the PPG signals, considering both physiological and demographical (age and gender) data. A Recursive Feature Elimination (RFE) algorithm was employed to choose the most contributive features for Hb prediction. An ensemble model combining several independent Extreme Learning Machine (ELM) models was established to enhance prediction stability and accuracy. A RMSE value of 1.72 and a PCC value of 0.76, indicating a strong correlation between predicted and actual Hb values, were achieved at the end of the experiments. Additionally, the proposed model outperformed other regression models (LR, SVR, RF, and traditional ELM) in terms of RMSE and PCC. The study introduced an ensemble approach to the ELM algorithm for improved prediction accuracy and stability, showcasing potential for broader clinical application and research into non-invasive biomarker detection.

A study, incorporates deep neural semantic segmentation and convolutional neural networks (CNNs), was presented by Chen et al. [20]. The study involved images of 1065 patients undergoing surgery. Hemoglobin levels among these patients ranged from 6 to 18 mol/L. The dataset was balanced by using the SMOTE algorithm due

to the original imbalance, where most patients had normal Hb levels. Deep neural segmentation was utilized to identify the palpebral conjunctiva region from images, ensuring the focus on relevant features for Hb prediction. CNNs and an ensemble of ELM were employed to predict Hb values. The proposed model obtained an  $R^2$  value of 0.512, indicating a strong predictive capability. The explained variance score (EVS) reached 0.535, and MAE was 1.521, demonstrating the accuracy of the prediction model. Compared to other methods like decision trees (DT), LR, and SVR, the suggested approach demonstrated enhanced performance in terms of  $R^2$ , EVS, and MAE. The research demonstrated the potential of using deep learning and image analysis for non-invasive hemoglobin level prediction. Additionally, the study highlighted the importance of causal knowledge in improving prediction accuracy and reducing the impact of pseudo-correlation noise in the images.

Kwon and Kim proposed a non-invasive method for estimating glycosylated hemoglobin (HbA1c) levels using PPG signals [21]. Their dataset was derived from 40 volunteers, including their PPG signals and corresponding HbA1c levels, measured invasively. Additional data such as body mass index (BMI), finger width (FW), and SpO2 levels were collected. For the experiments, a custom-developed device that measures PPG signals through both reflective and transmissive methods was utilized. 18 features were initially considered, based on physiological characteristics, signal-directed characteristics, and physical parameters. 7 key features were ultimately selected for their importance in estimating HbA1c levels, including zero-crossing rate (ZCR), power spectral density (PSD) variance, and FW. For feature selection, RFE and importance analysis were employed to identify the most contributive features for HbA1c prediction. The study utilized RF and XGBoost ensemble models for the prediction of HbA1c levels based on the extracted PPG signal features. XGBoost model showed superior performance with a PCC value of 0.957 for the reflection method including FW as a feature. For diabetes classification, XGBoost also outperformed RF, Beer-Lambert Model, and Photon-Diffusion Model. The study demonstrated that XGBoost model can provide a promising tool for diabetes management without the need for invasive blood samples.

### 3. MATERIAL AND METHODS

**3.1. Dataset.** The dataset utilized in this study comprises PPG signals, gender, age, and Hb value, designed for research into non-invasive hemoglobin measurement using machine learning [12]. The dataset includes 68 participants (56% female, 44% male) between the ages of 18 and 65. A total of 816 data points were collected, corresponding to 12 data points per participant. Red and infra-red light intensity values, measured by the PPG sensor in arbitrary units (a.u.), are represented as

numerical data (float). Gender corresponds to the categorical variable indicating each participant's sex. Age corresponds to the numerical variable (integer), indicating the participant's age in years. Hemoglobin (Hb) corresponds to the target variable indicating the concentration of hemoglobin in blood, measured in grams per deciliter (g/dL) as a numeric data type (float). Hb values were measured invasively. For each participant, raw PPG signals were gathered every 40 milliseconds across a span of 10 seconds. These signals were then averaged into 12 sets of red and infrared data to create the dataset. The study acknowledges potential biases in PPG signal measurements due to vibrations, movements, and subjects' skin tones, and the dataset's generalizability to different populations. The dataset is publicly accessible for further research and development in the field of non-invasive hemoglobin measurement and is hosted on Mendeley Data.

**3.2. General Framework.** The first stage in the study is to prepare the features to feed the machine learning algorithms. The raw signals, time-domain features, and frequency-domain features are utilized to build different models. Time-domain and frequency-domain features are extracted from the raw signals. Additionally, normalization is applied to improve the model performance and have interpretable results, while feature selection and feature reduction algorithms are applied to decrease the dimension of the samples. The second stage is to train the machine learning algorithms. In this stage, the aim is to optimize the hyper-parameters for a machine learning algorithm. Cross-validation technique is applied on the training data to obtain the performance values for the related hyper-parameters then this information is utilized by the Bayesian optimization algorithm to fine-tune the hyper-parameters. The final stage includes the test phase. After building the models, they are tested with the independent test set to obtain the performance results. The general framework is given in Figure 1.

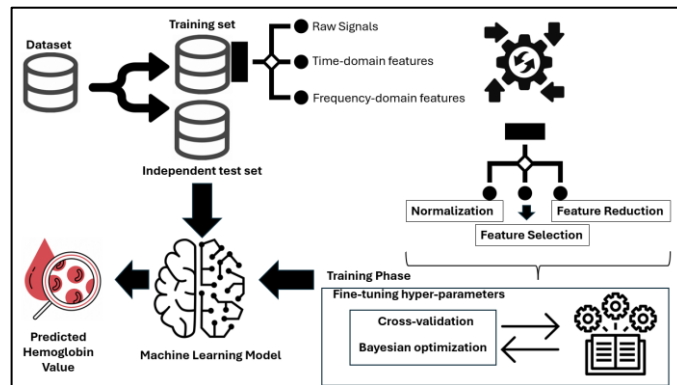


FIGURE 1. The general framework.

**3.3. Normalization.** Many machine learning algorithms perform better or require that the input data be normally distributed. Standardizing variables helps meet these assumptions or improve the algorithm's performance. Z-score normalization process, utilized in this study, involves transforming the original dataset so that the mean of the transformed data is 0 and the standard deviation is 1 for each feature. Z-score normalization is a valuable preprocessing step that can improve model performance. Its formula is given as follows:

$$z = \frac{(x-\mu)}{\sigma} \quad (1)$$

where,  $x$  represents the original value to be normalized,  $\mu$  represents the mean of the selected feature, and  $\sigma$  represents standard deviation of the same feature.

**3.4. Feature Extraction.** Time-domain and frequency-domain features were extracted from the raw signals to compare them and to feed machine learning regression algorithms. Time-domain features utilized in the study are mean, root mean square (RMS), standard deviation, shape factor, signal-to-noise ratio (SNR), signal to noise and distortion ratio (SINAD), peak value, crest factor, clearance factor and impulse factor. The set of frequency-domain features include mean frequency, median frequency, bandpower, occupied bandwidth power bandwidth, peak amplitude, peak location, and power spectral density (PSD) estimate.

**3.5. Feature Selection.** In our study, rReliefF feature selection algorithm was utilized. It is used primarily to identify relevant features that contribute significantly to the prediction of the output variable. Instead of looking for nearest neighbours within the same class or different classes, rReliefF for regression searches for  $k$  nearest neighbours based on the closeness of their response values [22]. The algorithm assesses how well a feature can discriminate between instances that are near each other in the feature space but have different response values. The process has 4 main steps. First, weights of each feature are set to 0. Second, an instance is selected randomly from the dataset. Third, for the selected instance, a set of nearest neighbours is found based on the feature space. Finally, for each feature, its weight is updated based on how much the feature values for selected instance and its nearest neighbours differ, considering the differences in their response values. The intuition is that if small differences in a feature correspond to large differences in the response variable for otherwise similar instances, then the feature is important for predicting the response.

**3.6. Feature Reduction.** In our study, Principal Component Analysis (PCA) algorithm was utilized for dimensionality reduction. PCA works by identifying the



axes (principal components) that maximize the variance in a dataset [23]. These principal components are orthogonal to each other, which guarantees that they capture distinct aspects or patterns within the data. The first principal component captures the most variance, the second captures the second most, and so on, allowing for dimensionality reduction by selecting a subset of components to retain while minimizing information loss. The principal components serve as novel features that can be used in a regression model. These features represent linear combinations of the initial variables and are chosen since they explain the maximum amount of variance in the data.

### 3.7. Machine Learning Regression Algorithms.

*3.7.1. Support Vector Machines (SVM).* Support Vector Regression (SVR) extends the concept of SVM from classification to regression problems. It incorporates the core principles of SVMs to handle regression, providing a unique approach to predict continuous outcomes. Unlike traditional regression methods that aim to minimize the error between the predicted and actual outcomes, SVR focuses on ensuring that errors do not exceed a certain threshold [24]. This is achieved by fitting the best line or hyperplane within a predefined margin of tolerance, effectively capturing as many data points as possible while ignoring errors that are within the acceptable range. This guarantees that the model does not excessively react to minor fluctuations in the training data, leading to more stable and generalizable predictions.

In regression tasks, SVR is employed by choosing a type of kernel (linear, polynomial, or radial basis function) to transform the original data into a higher-dimensional space where a linear regression surface seems likely to fit better. The SVR model then focuses on minimizing the error for only those data points that fall outside the epsilon margin, ignoring errors within the margin. This approach allows the SVR to balance the intricacy of the model and the extent to which deviations exceeding epsilon are acceptable.

*3.7.2. Artificial Neural Networks (ANNs).* ANNs are a foundational element of machine learning and artificial intelligence, drawing inspiration from the human brain's architecture and operations. When applied to regression tasks, they're often referred to as neural network regressors. An ANN consists of interconnected processing units or nodes, called neurons. There are simply 3 different layers in an ANN. The input layer takes the features, the neurons in the hidden layers operate on the features, and the output layer produces the predicted value [25]. Every link between neurons has a corresponding weight, which is adjusted during the learning process. In a regression context, ANNs are designed to predict continuous outcomes based on input features, as opposed to classifying inputs into categories. The aim of

an ANN regressor is to learn a mapping from inputs to a continuous output, minimizing the disparity between forecasted and real values across a training dataset. This involves adjusting the weights of the connections in the network to minimize a loss function, a measure of prediction error, through a process known as backpropagation.

The model's complexity and capacity can be adjusted by varying the number of hidden layers and neurons within them, allowing ANNs to model complex, nonlinear relationships that might be difficult for other regression techniques to capture. However, they require careful tuning of hyperparameters and feature scaling, especially as model complexity increases.

*3.7.3. Classification and Regression Trees (CART).* Classification and Regression Trees (CART) is a decision tree learning technique that can be used for both classification and regression predictive modelling problems. The method involves splitting data into subsets based on the value of input features, leading to a tree-like model of decisions and their possible consequences [26]. The main goal of CART is to develop a model capable of predicting the value of a target variable by deriving straightforward decision rules from the features present in the data.

In regression tasks, CART involves building a decision tree to model the relationship between the features of data and a continuous target variable. The data is split at nodes based on feature values, aiming to minimize the variance of the target variable within each node. The process continues until a stopping criterion is met, like a maximum depth of the tree or a minimum number of samples in a node. The outcome is a model where each leaf node represents a prediction value based on the input features.

3.7.4. *Ensembles of Trees (EoT)*. Ensembles of trees are advanced machine learning techniques that combine multiple decision trees to create a more powerful model. These models are used for both classification and regression tasks. The core idea behind ensemble methods is to leverage the collective power of multiple models to achieve better accuracy and performance than any single model could on its own. Ensemble methods involve the integration of multiple decision trees to form a stronger predictor. Bootstrap Aggregating (Bagging) and Least Squares Boosting (LSBoost) can be given as examples for the most common ensemble methods [27,28]. LSBoost is a gradient boosting method that uses least squares loss to improve models' predictions iteratively. Bagging involves training multiple models in parallel, each on a random subset of the data (with replacement), and then aggregating their predictions. This approach is effective in reducing variance and overfitting. In EoT, a higher number of decision trees (learners) can increase the accuracy but may also lead to increased computational complexity and the risk of overfitting. The minimum leaf size in trees refers to the smallest number of observations that must be present in the leaf (terminal node) of a tree. Setting a higher minimum leaf size can help prevent overfitting by ensuring that the trees are not too deep or overly complex, which might make them sensitive to noise in the training data. By aggregating the predictions of multiple trees, ensembles can capture more complex patterns in the data, reduce the risk of overfitting, and handle variance better.

3.8. **Bayesian Optimization.** Bayesian optimization is a strategy used for optimizing objective functions that are expensive to evaluate [29]. It's particularly useful when dealing with black-box functions where the underlying mathematical form is unknown and derivatives are not available, making traditional optimization methods unsuitable. Bayesian optimization is widely used in machine learning and hyperparameter tuning where simulations or experiments are costly and time-consuming. Surrogate Model and Acquisition Function are crucial components in Bayesian optimization. They work together to efficiently find the minimum or maximum of an expensive function. Bayesian optimization builds a probabilistic model of the objective function, called the surrogate model, to approximate the true function. This model is used to make predictions about the function's behaviour and estimate the uncertainty of those predictions. Gaussian Processes (GP) are the most used surrogate models in Bayesian optimization owing to their capability to model the uncertainty of predictions. The acquisition function is used to decide where to sample next. It determines the trade-off between exploration (sampling where the model is uncertain) and exploitation (sampling where the model predicts high values). The acquisition function is chosen to be easily maximized unlike the original objective function. While Bayesian optimization can be used for regression, it does not directly target minimizing MSE or RMSE during the optimization process.

Instead, it focuses on finding the model parameters that have the highest posterior probability given the data and any prior beliefs. In other words, Bayesian optimization is a specific technique used to optimize expensive functions where the goal is to minimize or maximize the function's output. Here, the acquisition function within Bayesian optimization considers the uncertainty of the surrogate model to choose the next data point that will be most informative for finding the minimum or maximum.

One popular acquisition function is the Expected Improvement (EI), which measures the expected amount of improvement over the current best observation at a given point. The EI for a point  $x$  can be computed as follows while assuming minimization:

$$EI(x) = (\mu(x) - f(x^+) - \xi)\Phi(Z) + \sigma(x)\phi(Z) \quad (2)$$

$$Z = \begin{cases} \frac{\mu(x) - f(x^+) - \xi}{\sigma(x)}, & \sigma(x) > 0 \\ 0, & \sigma(x) \leq 0 \end{cases}$$

where,  $\mu(x)$  is the mean prediction of the surrogate model at  $x$ ,  $f(x^+)$  is the value of the best sample observed so far,  $\xi$  is a small positive number to encourage exploration,  $\sigma(x)$  is the standard deviation of the prediction at  $x$ ,  $\Phi$  and  $\phi$  represent the cumulative distribution function and probability density function of the standard normal distribution, respectively. The term  $Z$  is used to calculate the expected improvement. If the predictive uncertainty at  $x$  ( $\sigma(x)$ ) is zero, implying no uncertainty in the model's prediction at  $x$ ,  $Z$  is set to 0 since the formula aims to prevent division by zero.  $Z$  plays a crucial role in quantifying how much improvement a new sample is expected to provide over the current best observation, adjusted for the level of uncertainty in the prediction at that point. This standardization allows the EI formula to balance exploration and exploitation by taking into account both the average prediction and the uncertainty of the prediction.

When evaluating the objective function is time-consuming, it's beneficial to incorporate the evaluation time into the acquisition function. The Expected Improvement Per Second Plus (EIPS) is a variant of the EI that accounts for the evaluation time, aiming to maximize the efficiency of the optimization process in terms of the improvement gained per unit of time. The EIPS acquisition function can be formulated as:

$$EIPS(x) = \frac{EI(x)}{t(x)} \quad (3)$$

where,  $E(x)$  is the expected improvement at point  $x$  and  $t(x)$  is the expected

evaluation time for point  $x$ . This formulation encourages selecting points that are not only expected to yield high improvement but also are quicker to evaluate, thus optimizing the efficiency of the Bayesian optimization process [30].

#### 4. RESULTS

**4.1. Experimental Setup.** The dataset utilized in this study was split into training and independent test sets. The training set comprised 70% of the data (48 samples), while the independent test set included the remaining 30% (20 samples).

Given the relatively small size of our dataset, Leave-One-Out (LOO) cross-validation is a suitable choice for optimizing the hyperparameters of our machine learning algorithms during the training phase. In LOO strategy, each sample is used once as a validation case, while the remaining part of the training set is used to obtain a model. This process is repeated for every sample in the training set. Finally, by averaging the performances of all validation samples, hyper-parameters of a machine learning algorithm are determined. With the integration of Bayesian optimization algorithm, hyper-parameters are optimized, leading to improved model performance.

5 setups were prepared utilizing the same machine learning algorithms but with different features, in order to compare the performances in terms of evaluation metrics.

For each machine learning algorithm, the hyper-parameters were fine-tuned by utilizing Bayesian optimization. These hyper-parameters were box constraint (cost), epsilon, and kernel function (linear, Gaussian, quadratic, and cubic) regarding SVR model; number of hidden layers, size of each layer, activation function (sigmoid, rectified layer unit), and regularization strength for ANN model; minimum leaf size for CART model; ensemble method (LSBoost or Bag), number of learners, learning rate, minimum leaf size, and number of features to sample for EoT model.

**4.2. Evaluation Metrics.** Mean Absolute Error (MAE) quantifies the average size of the mistakes in a series of forecasts, disregarding their sign. It calculates the average of the absolute differences between the forecasted and the actual values. Below is the formula for MAE:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

where  $n$ ,  $y_i$ , and  $\hat{y}_i$  represent the number of observations, the actual value of the observation, and the predicted value, respectively. The lower the MAE, the better, with 0 being the ideal score.

Mean Squared Error (MSE) evaluates the average of the squared discrepancies, differences between the predicted values and the true values. The formula for MSE is provided below:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

where,  $n$ ,  $y_i$ , and  $\hat{y}_i$  represent the same meanings in Equation (4). It penalizes larger errors more severely than smaller ones, due to the squaring of each term. A smaller MSE signifies a closer match to the actual data, where a score of 0 represents an ideal fit.

Root Mean Squared Error (RMSE) calculates the square root of the mean of the squared deviations between the predicted values and the actual observations. It provides an indication of the dispersion of these residuals, essentially showing the degree to which the data clusters around the best fit line. Below is the formula for RMSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

where,  $n$ ,  $y_i$ , and  $\hat{y}_i$  represent the same meanings in Equation (5), but the whole formula is under a square root. RMSE is the square root of MSE, bringing the error metric back to the same units as the target variable. It similarly penalizes larger errors more than smaller ones. smaller RMSE value suggests a more accurate model, with 0 being the ideal score. RMSE is sensitive to outliers.

R-squared ( $R^2$ ) quantifies the fraction of variance in the dependent variable that can be explained by the independent variables. The formula for  $R^2$  is as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

where,  $n$ ,  $y_i$ , and  $\hat{y}_i$  represent the same meanings in Equation (4),  $\bar{y}$  represents the mean of the actual values.  $SS_{res}$  is the sum of squares of residuals, which measures the variability of the prediction errors.  $SS_{tot}$  is the total sum of squares, which measures the total variability of the observed data around the mean. The nearer  $R^2$  approaches 1, the greater the proportion of variance in the dependent variable explained by the model, signifying a stronger model fit. An elevated  $R^2$  value does not automatically mean the model is the most effective or accurate in its predictions. In models where the predictions are worse than merely estimating the average of the observed values,  $R^2$  can be negative.

4.3. **Empirical Results.** The results of the first, second, third, fourth, and fifth setups are given in Table 1, Table 2, Table 3, Table 4, and Table 5, respectively.

TABLE 1. Regression results for raw signals, age and gender.

Raw signals + Age + Gender (26 features)								
Validation					Test			
Model	MAE	MSE	RMSE	R <sup>2</sup>	MAE	MSE	RMSE	R <sup>2</sup>
SVR	1.419	3.152	1.775	0.157	<b>1.016</b>	<b>1.783</b>	<b>1.335</b>	<b>0.085</b>
ANN	1.232	2.111	1.453	0.435	1.579	4.1	2.024	-1.01
CART	1.099	2.052	1.432	0.451	1.254	2.96	1.721	-0.518
EoT	<b>0.913</b>	<b>1.732</b>	<b>1.316</b>	<b>0.537</b>	1.17	2.529	1.591	-0.297
RReliefF (19 features)								
SVR	1.434	2.939	1.714	0.214	<b>1.044</b>	<b>1.859</b>	<b>1.363</b>	<b>0.046</b>
ANN	<b>1.075</b>	<b>1.729</b>	<b>1.315</b>	<b>0.537</b>	1.677	4.523	2.126	-1.32
CART	1.12	2.099	1.448	0.439	1.331	2.972	1.724	-0.525
EoT	1.147	2.712	1.646	0.275	1.422	3.769	1.941	-0.933
PCA (95% variance, 2 components)								
SVR	<b>1.506</b>	<b>3.320</b>	<b>1.822</b>	<b>0.112</b>	1.247	2.443	1.563	-0.253
ANN	1.532	3.562	1.887	0.047	<b>1.101</b>	<b>2.149</b>	<b>1.466</b>	<b>-0.102</b>
CART	1.52	3.693	1.921	0.012	1.293	3.035	1.742	-0.557
EoT	1.63	3.804	1.95	-0.017	1.225	2.799	1.673	-0.436
RReliefF + PCA								
SVR	<b>1.504</b>	<b>3.314</b>	<b>1.82</b>	<b>0.113</b>	1.24	2.403	1.55	-0.233
ANN	1.638	4.052	2.013	-0.083	<b>1.131</b>	<b>2.026</b>	<b>1.423</b>	<b>-0.039</b>
CART	1.576	3.58	1.892	0.042	1.307	2.526	1.589	-0.296
EoT	1.576	3.739	1.933	4.4e-16	1.219	2.289	1.513	-0.174

According to the Table 1, the best performance values were obtained as 1.016, 1.783, 1.335, and 0.085 in terms of MAE, MSE, RMSE, and R<sup>2</sup>, respectively for the independent test set by utilizing raw signals, age, and gender information as features and SVR as regressor.

TABLE 2. Regression results for normalized raw signals, age and gender.

Model	Normalized features (26 features)							
	Validation				Test			
	MAE	MSE	RMSE	R <sup>2</sup>	MAE	MSE	RMSE	R <sup>2</sup>
SVR	1.143	2.213	1.487	0.408	1.432	3.325	1.823	-0.705
ANN	<b>1.015</b>	<b>1.508</b>	<b>1.228</b>	<b>0.597</b>	1.718	4.338	2.082	-1.225
CART	1.099	2.052	1.432	0.451	1.254	2.96	1.72	-0.518
EoT	1.22	2.585	1.608	0.308	<b>1.083</b>	<b>2.314</b>	<b>1.521</b>	<b>-0.187</b>
RRReliefF (19 features)								
SVR	1.132	1.991	1.411	0.467	<b>0.943</b>	<b>1.764</b>	<b>1.328</b>	<b>0.095</b>
ANN	1.092	2.016	1.42	0.461	1.136	1.806	1.344	0.073
CART	<b>1.054</b>	<b>1.937</b>	<b>1.392</b>	<b>0.482</b>	1.254	2.96	1.72	-0.518
EoT	<b>1.054</b>	2.095	1.447	0.439	1.277	2.76	1.661	-0.416
PCA (95% variance, 3 components)								
SVR	<b>0.903</b>	<b>1.218</b>	<b>1.103</b>	<b>0.674</b>	1.706	4.605	2.146	-1.362
ANN	1.091	2.234	1.494	0.403	<b>1.000</b>	<b>1.556</b>	<b>1.247</b>	<b>0.202</b>
CART	1.122	2.063	1.436	0.448	1.116	2.178	1.475	-0.117
EoT	1.227	2.292	1.514	0.387	1.022	2.038	1.427	-0.045
RRReliefF + PCA								
SVR	<b>0.96</b>	<b>1.4</b>	<b>1.184</b>	<b>0.625</b>	2.845	11.999	3.464	-5.155
ANN	1.042	2.152	1.467	0.424	<b>0.981</b>	<b>1.508</b>	<b>1.228</b>	<b>0.226</b>
CART	1.226	2.257	1.502	0.396	1.097	2.036	1.426	-0.044
EoT	1.255	2.508	1.583	0.329	1.072	2.163	1.471	-0.11

According to the Table 2, the best MAE value was obtained as 0.943 by utilizing SVR and feature selection. In terms of MSE, RMSE, and R<sup>2</sup>, the best performance values were achieved as 1.508, 1.228, and 0.226, respectively for the independent test set by utilizing normalized features, ANN as regressor, and with the inclusion of the feature selection and reduction methods.



TABLE 3. Regression results for normalized raw signals and w/o age and gender.

Normalized features (24 features)								
Model	Validation				Test			
	MAE	MSE	RMSE	R <sup>2</sup>	MAE	MSE	RMSE	R <sup>2</sup>
SVR	1.463	3.34	1.827	0.107	<b>1.004</b>	<b>1.951</b>	<b>1.396</b>	<b>-0.001</b>
ANN	<b>1.295</b>	<b>2.572</b>	<b>1.603</b>	<b>0.312</b>	1.267	2.416	1.554	-0.239
CART	1.576	3.548	1.883	0.051	1.181	2.416	1.555	-0.239
EoT	1.677	3.817	1.953	-0.021	1.184	2.598	1.612	-0.333
RReliefF (16 features)								
SVR	1.498	3.279	1.811	0.123	1.193	2.392	1.546	-0.227
ANN	<b>1.397</b>	<b>3.277</b>	<b>1.81</b>	<b>0.124</b>	<b>1.171</b>	<b>2.179</b>	<b>1.476</b>	<b>-0.117</b>
CART	1.575	3.547	1.883	0.051	1.181	2.416	1.554	-0.239
EoT	1.645	3.957	1.989	-0.058	1.225	2.482	1.575	-0.273
PCA (95% variance, 2 components)								
SVR	<b>1.501</b>	<b>3.304</b>	<b>1.817</b>	<b>0.116</b>	1.236	2.397	1.548	-0.23
ANN	1.602	3.821	1.954	-0.021	1.256	2.45	1.565	-0.256
CART	1.576	3.739	1.933	0	1.258	2.39	1.546	-0.226
EoT	1.59	3.687	1.92	0.013	<b>1.182</b>	<b>2.233</b>	<b>1.494</b>	<b>-0.146</b>
RReliefF + PCA								
SVR	<b>1.523</b>	<b>3.31</b>	<b>1.819</b>	<b>0.115</b>	1.23	2.401	1.549	-0.232
ANN	1.563	3.768	1.941	-0.007	<b>1.16</b>	<b>2.076</b>	<b>1.442</b>	<b>-0.066</b>
CART	1.602	3.547	1.883	0.051	1.361	3.143	1.772	-0.612
EoT	1.638	3.907	1.976	-0.044	1.259	2.451	1.565	-0.257

According to the Table 3, the best performance values were obtained as 1.004, 1.951, 1.396, and -0.001 in terms of MAE, MSE, RMSE, and R<sup>2</sup>, respectively, by utilizing SVR and the normalized features without the gender and the age information for the independent test set.

TABLE 4. Regression results for normalized time-domain features.

Normalized time-domain features (10 features)								
Model	Validation				Test			
	MAE	MSE	RMSE	R <sup>2</sup>	MAE	MSE	RMSE	R <sup>2</sup>
SVR	<b>1.434</b>	<b>2.819</b>	<b>1.679</b>	<b>0.246</b>	<b>1.183</b>	2.671	1.634	-0.371
ANN	1.51	3.151	1.775	0.157	1.238	<b>2.595</b>	<b>1.611</b>	<b>-0.331</b>
CART	1.519	3.233	1.798	0.135	1.199	2.608	1.615	-0.338
EoT	1.612	3.662	1.913	0.02	1.232	3.354	1.831	-0.721
RReliefF (4 features)								
SVR	1.539	3.415	1.848	0.086	1.111	<b>2.089</b>	<b>1.445</b>	<b>-0.071</b>
ANN	<b>1.532</b>	<b>3.279</b>	<b>1.811</b>	<b>0.123</b>	<b>1.09</b>	2.124	1.457	-0.089
CART	1.587	3.492	1.868	0.066	1.199	2.608	1.615	-0.338
EoT	1.59	3.526	1.877	0.056	1.267	3.288	1.813	-0.686
PCA (95% variance, 1 component)								
SVR	1.549	3.68	1.918	0.015	1.159	2.077	1.441	-0.065
ANN	1.521	3.671	1.916	0.018	<b>1.056</b>	1.968	1.402	-0.009
CART	<b>1.51</b>	<b>3.616</b>	<b>1.901</b>	<b>0.032</b>	1.269	2.258	1.502	-0.158
EoT	1.561	3.924	1.981	-0.049	1.145	<b>1.726</b>	<b>1.313</b>	<b>0.114</b>
RReliefF + PCA (2 components)								
SVR	<b>1.358</b>	<b>3.116</b>	<b>1.765</b>	<b>0.166</b>	1.319	2.889	1.699	-0.482
ANN	1.575	3.572	1.89	0.044	<b>1.147</b>	<b>2.352</b>	<b>1.533</b>	<b>-0.206</b>
CART	1.562	3.61	1.9	0.034	1.291	2.723	1.65	-0.397
EoT	1.624	3.863	1.965	-0.033	1.236	2.89	1.7	-0.482

According to the Table 4, the best MAE value was obtained as 1.056 on normalized time-domain features, extracted from the raw signals, by utilizing ANN as regressor and the feature reduction method. The best performance values in terms of MSE, RMSE, and R<sup>2</sup> were achieved as 1.726, 1.313, and 0.114, respectively by utilizing EoT as regressor and the feature reduction method.

TABLE 5. Regression results for normalized frequency-domain features.

Normalized frequency-domain features (136 features)								
Model	Validation				Test			
	MAE	MSE	RMSE	R <sup>2</sup>	MAE	MSE	RMSE	R <sup>2</sup>
SVR	1.574	3.762	1.939	-0.006	1.155	<b>2.061</b>	<b>1.435</b>	<b>-0.057</b>
ANN	1.571	3.404	1.845	0.089	1.246	2.983	1.727	-0.53
CART	<b>1.493</b>	<b>3.368</b>	<b>1.835</b>	<b>0.099</b>	1.282	2.853	1.689	-0.463
EoT	1.536	3.389	1.841	0.093	<b>1.125</b>	2.195	1.481	-0.126
RReliefF (2 features)								
SVR	<b>1.426</b>	<b>3.107</b>	<b>1.762</b>	<b>0.168</b>	1.241	2.734	1.653	-0.402
ANN	1.629	3.917	1.979	-0.047	<b>1.205</b>	<b>2.237</b>	<b>1.495</b>	<b>-0.147</b>
CART	1.515	3.231	1.797	0.135	1.397	3.759	1.938	-0.928
EoT	1.519	3.278	1.81	0.123	1.227	2.689	1.639	-0.379
PCA (95% variance, 2 components)								
SVR	<b>1.451</b>	<b>3.212</b>	<b>1.792</b>	<b>0.14</b>	1.446	3.081	1.755	-0.58
ANN	1.62	3.933	1.983	-0.052	1.48	3.542	1.882	-0.817
CART	1.535	3.252	1.803	0.13	<b>1.037</b>	2.146	1.465	-0.101
EoT	1.67	3.854	1.963	-0.031	1.082	<b>2.113</b>	<b>1.453</b>	<b>-0.084</b>
RReliefF + PCA (2 components)								
SVR	1.488	3.159	1.777	0.155	1.373	3.348	1.83	-0.718
ANN	<b>1.441</b>	<b>2.959</b>	<b>1.72</b>	<b>0.208</b>	1.456	3.723	1.929	-0.91
CART	1.468	3.384	1.839	0.094	<b>1.251</b>	<b>2.598</b>	<b>1.611</b>	<b>-0.332</b>
EoT	1.598	3.636	1.907	0.027	1.265	2.993	1.73	-0.535

According to the Table 5, the best MAE value was obtained as 1.037 on normalized frequency-domain features, extracted from the raw signals, by utilizing CART as regressor and the feature reduction method. The other best performance values were achieved as 2.061, 1.435, and -0.057 in terms of MSE, RMSE, and R<sup>2</sup>, respectively, by utilizing SVR without the inclusion of the feature selection and reduction methods.

The best performance values, according to the independent test set results, are shown in Table 6, for each experimental setup.

TABLE 6. The best regression results for 5 experimental setups.

Setup 1: Raw signals + Age + Gender (26 features)								
Model	Validation				Test			
	MAE	MSE	RMSE	R <sup>2</sup>	MAE	MSE	RMSE	R <sup>2</sup>
SVR	1.419	3.152	1.775	0.157	<b>1.016</b>	<b>1.783</b>	<b>1.335</b>	<b>0.085</b>
Setup 2: Normalized features + RReliefF (19 features)								
SVR	1.132	1.991	1.411	0.467	<b>0.943</b>	1.764	1.328	0.095
Setup 2: Normalized features + RReliefF + PCA (3 components)								
ANN	1.042	2.152	1.467	0.424	0.981	<b>1.508</b>	<b>1.228</b>	<b>0.226</b>
Setup 3: Normalized features w/o Age & Gender (24 features)								
SVR	1.463	3.34	1.827	0.107	<b>1.004</b>	<b>1.951</b>	<b>1.396</b>	<b>-0.001</b>
Setup 4: Normalized time-domain features + PCA (1 component)								
ANN	1.521	3.671	1.916	0.018	<b>1.056</b>	1.968	1.402	-0.009
EoT	1.561	3.924	1.981	-0.049	1.145	<b>1.726</b>	<b>1.313</b>	<b>0.114</b>
Setup 5: Normalized frequency-domain features + PCA (2 components)								
CART	1.535	3.252	1.803	0.13	<b>1.037</b>	2.146	1.465	-0.101
Setup 5: Normalized frequency-domain features (136 features)								
SVR	1.574	3.762	1.939	-0.006	1.155	<b>2.061</b>	<b>1.435</b>	<b>-0.057</b>

According to the Table 6, for the first experimental setup, the best performance values were obtained by SVR regressor on raw signals with the inclusion of age and gender information. For the second experimental setup, the best performance value was obtained by SVR regressor on normalized features with the inclusion of the feature selection algorithm in terms of MAE, while the best performance values were obtained by ANN regressor with the inclusion of the feature selection and reduction algorithms in terms of MSE, RMSE, and R<sup>2</sup>. For the third experimental setup, the best performance values were obtained by SVR regressor on normalized features without the inclusion of age and gender information. For the fourth experimental setup, the best performance value was obtained by ANN regressor on normalized time-domain features with the inclusion of the feature reduction algorithm in terms of MAE, while the best performance values were obtained by EoT regressor in terms of the other evaluation metrics. For the last experimental setup, the best performance value was obtained by CART regressor on frequency-domain features with the inclusion of the feature reduction algorithm in terms of MAE, while the best performance values were obtained by SVR regressor on frequency-domain features without the inclusion of the feature selection and reduction algorithms in terms of the other evaluation metrics.

In the first setup, for the SVR model, the hyperparameters were optimized by Bayesian optimization with a cost of 0.279, epsilon value of 1.116, and a kernel function of linear. In the second setup, for the SVR model, the hyperparameters were optimized with a cost of 0.207, epsilon value of 0.364, and a kernel function of

quadratic, while for the ANN model the number of hidden layers, the number of neurons, activation function, and regularization strength were optimized as 1, 1, sigmoid, and 1.385, respectively. In the third setup, for the SVR model, the hyperparameters were optimized with a cost of 0.001, epsilon value of 0.002, and a kernel function of quadratic. In the fourth setup, for the ANN model, the number of hidden layers, the number of neurons, activation function, and regularization strength were optimized as 1, 1, rectified layer unit, and 0.397, respectively. For the EoT model, the ensemble method, the number of learners, the learning rate, the minimum leaf size, and the number of the features to sample were optimized as LSBoost, 24, 0.998, 15, and 1, respectively. In the last setup, the minimum leaf size was optimized as 19 for the CART model, while the hyperparameters were optimized with a cost of 0.001, epsilon value of 0.013, and a kernel function of Gaussian, for the SVR model.

## 5. CONCLUSION

According to the data obtained from the experiments, it was observed that standardization on the features is an important preprocessing step. Experiments further showed that age and gender were informative features, as the performance of the regressors dropped when these features were removed from the feature vectors. When time-domain and frequency-domain features were used to feed the regressors without including age and gender information, the results showed that the time-domain features led to better performance than the frequency-domain-features. Based on  $R^2$  evaluation metric, normalized raw signal and age-gender information can explain the dependent variable better than other features. It can be interpreted that the utilization of time and frequency-domain features is indicative of potential information loss relative to the raw signal. If we compare the models, it can be observed that while the SVR model stood out in 4 out of 5 different setups, the highest performance values were obtained with the ANN model. ANN provided the best performance values among the setups in terms of MSE, RMSE, and  $R^2$  after normalization, feature selection, and reduction were applied. We hope the results obtained from the first utilization of this relevant dataset will be established as a benchmark, encouraging further research, and paving the way for achieving even better results in the future.

**Declaration of Competing Interests** The author declares no known competing interests.

## REFERENCES

- [1] Cho, H., Lee, S.-R., Baek, Y., Anemia diagnostic system based on impedance measurement of red blood cells, *Sensors*, 21 (23) (2021), 1-12, <https://doi.org/10.3390/s21238043>.
- [2] Yap, B. K., Soair, S. N. M., Talik, N. A., Lim, W. F., Mei, I. L., Potential point-of-care microfluidic devices to diagnose iron deficiency anemia, *Sensors*, 18 (8) (2018), 1-17, <https://doi.org/10.3390/s18082625>.
- [3] Mandal, A. K., Mitra, A., Das, R., Sickle cell hemoglobin, In: Hoeger, U., Harris, J. (eds), Vertebrate and invertebrate respiratory proteins, lipoproteins and other body fluid proteins, *Subcellular Biochemistry*, vol 94, Springer, Cham, (2020), [https://doi.org/10.1007/978-3-030-41769-7\\_12](https://doi.org/10.1007/978-3-030-41769-7_12).
- [4] Telfer, P., Carvalho, S. J., Ruzangi, J., Arici, M., Binns, M., Beaubrun, A., Montealegre-Golcher, F., Rice, C. T., Were, J. J., Association between hemoglobin levels and end-organ damage in sickle cell disease: A retrospective linked primary and secondary care database analysis in England, *Hematol. Transfus. Cell Ther.*, 44 (Supplement 2) (2022), S10-S11.
- [5] Helmi, N., Bashir, M., Shireen, A., Ahmed, I. M., Thalassemia review: features, dental considerations and management, *Electron. Physician*, 9 (3) (2017), 4003-4008, <https://doi.org/10.19082/4003>.
- [6] Gaspar, B. L., Sharma, P., Das, R., Anemia in malignancies: Pathogenetic and diagnostic considerations, *Hematology*, 20 (1) (2015), 18-25, <https://doi.org/10.1179/1607845414Y.0000000161>.
- [7] Panjeta, M., Tahirović, I., Sofić, E., Ćorić, J., Dervišević, A., Interpretation of erythropoietin and haemoglobin levels in patients with various stages of chronic kidney disease, *J. Med. Biochem.*, 36 (2) (2017), 145-152, <https://doi.org/10.1515/jomb-2017-0014>.
- [8] World Health Organization, Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity, (2011). Available at: <https://www.who.int/publications/i/item/WHO-NMH-NHD-MNM-11.1>. [Accessed March 2024].
- [9] Hasan, M. K., Aziz, M. H., Zarif, M. I. I., Hasan, M., Hashem, M., Guha, S., Love, R. R., Ahamed, S., Noninvasive hemoglobin level prediction in a mobile phone environment: State of the art review and recommendations, *JMIR mHealth and uHealth*, 9 (4) (2021), 1-24, <https://doi.org/10.2196/16806>.
- [10] Peng, F., Zhang, N., Chen, C., Wu, F., Wang, W., Ensemble extreme learning machine method for hemoglobin estimation based on photoplethysmographic signals, *Sensors*, 24 (6) (2024), 1-14, <https://doi.org/10.3390/s24061736>.
- [11] Zhu, J., Sun, R., Liu, H., Wang, T., Cai, L., Chen, Z., Heng, B., A non-invasive hemoglobin detection device based on multispectral photoplethysmography, *Biosensors*, 14 (1) (2024), 1-19, <https://doi.org/10.3390/bios14010022>.

- [12] Abuzairi, T., Vinia, E., Yudhistira, M. A., Rizkinia, M., Eriska, W., A dataset of hemoglobin blood value and photoplethysmography signal for machine learning-based non-invasive hemoglobin measurement, *Data in Brief*, 52 (2024), 1-7, <https://doi.org/10.1016/j.dib.2023.109823>.
- [13] Dimauro, G., Caivano, D., Girardi, F., A new method and a non-invasive device to estimate anemia based on digital images of the conjunctiva, *IEEE Access*, 6 (2018), 1-8, <https://doi.org/10.1109/ACCESS.2018.2867110>.
- [14] Ding, H., Lu, Q., Gao, H., Peng, Z., Non-invasive prediction of hemoglobin levels by principal component and back propagation artificial neural network, *Biomed. Opt. Express*, 5 (2014), 1145-1152, <https://doi.org/10.1364/BOE.5.001145>.
- [15] Wang, E. J., Li, W., Zhu, J., Rana, R., Patel, S. N., Noninvasive hemoglobin measurement using unmodified smartphone camera and white flash, *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju, Korea (South), (2017), 2333-2336, <https://doi.org/10.1109/EMBC.2017.8037323>.
- [16] Kavsaoglu, A., Polat, K., Hariharan, M., Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristic features, *Appl. Soft Comput.*, 28 (2015), 433-441, <https://doi.org/10.1016/j.asoc.2015.04.008>.
- [17] Hasan, M. K., Haque, M. M., Adib, R., Tumpa, J. F., Begum, A., Love, R. R., Kim, Y. L., Sheikh, I. A., SmartHeLP: Smartphone-based hemoglobin level prediction using an artificial neural network, *AMIA Annu. Symp. Proc.*, 2018, 535-544.
- [18] El-kenawy, E. S. M. T., A machine learning model for hemoglobin estimation and anemia classification, *IJCSIS*, 17 (2) (2019), 100-108.
- [19] Chen, Z., Qin, H., Ge, W., Li, S., Liang, Y., Research on a non-invasive hemoglobin measurement system based on four-wavelength photoplethysmography, *Electronics*, 12 (6) (2023), 1-12, <https://doi.org/10.3390/electronics12061346>.
- [20] Chen, Y., Zhong, K., Zhu, Y., Sun, Q., Two-stage hemoglobin prediction based on prior causality, *Front. Public Health*, 10, (2022), 1-12, <https://doi.org/10.3389/fpubh.2022.1079389>.
- [21] Kwon, T.-H., Kim, K.-D., Machine-learning-based noninvasive in vivo estimation of HbA1c using photoplethysmography signals, *Sensors*, 22 (8) (2022), 1-19, <https://doi.org/10.3390/s22082963>.
- [22] Robnik-Sikonja, M., Kononenko, I., Theoretical and empirical analysis of ReliefF and RReliefF, *Mach. Learn.*, 53 (2003), 23-69.
- [23] Greenacre, M., Groenen, P. J. F., Hastie, T. et al., Principal component analysis, *Nat. Rev. Methods Primers*, 2 (2022), 100, <https://doi.org/10.1038/s43586-022-00184-w>.
- [24] Wang, Y. G., Wu, J., Hu, Z. H., McLachlan, G. J., A new algorithm for support vector regression with automatic selection of hyperparameters, *Pattern Recognit.*, 133 (2023), 1-9, <https://doi.org/10.1016/j.patcog.2022.108989>.
- [25] Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., et al., What is machine learning, artificial

- neural networks and deep learning? Examples of practical applications in medicine, *Diagnostics*, 13 (15) (2023), 1-22, <https://doi.org/10.3390/diagnostics13152582>.
- [26] Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J., *Classification and Regression Trees*, Boca Raton, FL: Chapman and Hall, 1984.
- [27] Breiman, L., Bagging predictors, *Mach. Learn.*, 26 (1996), 123-140.
- [28] Friedman, J., Hastie, T., Tibshirani, R., Additive logistic regression: A statistical view of boosting, *Ann. Stat.*, 28 (2) (2000), 337-407.
- [29] Wang, X., Jin, Y., Schmitt, S., Olhofer, M., Recent advances in Bayesian optimization, *ACM Comput. Surv.*, 55 (13s) (2023), 1-36, <https://doi.org/10.1145/3582078>.
- [30] Jones, D. R., Schonlau, M., Welch, W. J., Efficient global optimization of expensive black-box functions, *J. Glob. Optim.*, 13 (1998), 455-492.



## OPTICAL FIBER BENDING SENSOR BASED ON SPECKLE PATTERN IMAGING

İsa NAVRUZ<sup>1</sup>, Ceren DILSİZ<sup>1</sup>, Eylül Sevim ORTAK<sup>1</sup>, Sevde Nur BOYRAZ<sup>1</sup>





<sup>1</sup> Department of Electrical and Electronics Engineering, Ankara University,  
Ankara, TÜRKİYE

**ABSTRACT.** In this paper, we propose a new fiber bending sensor based on speckle pattern imaging. The design and implementation of the sensor are demonstrated by simulated studies. The speckle pattern imaging technique by using a multimode fiber can offer high spatial resolution. In this study, we showed that the bending sensor responds very sensitively by using the correlation of the images. The fiber sensing part consists of a curve in a form similar to the S structure. We reached a sensitivity of  $0.0295 \mu\text{m}^{-1}$  by bending the fiber only  $60^\circ$ . Sensitivity can be further increased by reducing the bending diameter or creating a full loop.

### 1. INTRODUCTION

Fiber optic sensors have emerged as a powerful and versatile technology for measuring various physical parameters with high accuracy and sensitivity. They use the basic principles of light transmission and modulation within optical fibers to measure. Unlike conventional sensors that rely on electrical or mechanical components, fiber optic sensors exploit the interaction between light and the surrounding environment to detect and quantify changes in temperature [1,2 3, 4], strain [5, 6, 7], pressure [8, 9, 10], humidity [11, 12, 13, 14], chemical composition [15, 16, 17, 18] and more. This ability to convert physical quantities into measurable optical signals has opened up new avenues for highly sensitive and reliable sensing. Fiber optic sensors offer numerous advantages such as small size, immunity to electromagnetic interference, remote sensing capability, and the ability to multiplex

*Keywords.* Speckle pattern, fiber sensor, bending sensor, multimode fiber, image correlation.

✉ inavruz@ankara.edu.tr-Corresponding author;  0000-0003-2976-076X  
✉ cerendilsiz260@gmail.com;  0009-0001-6486-0574  
✉ eylulortak006@gmail.com;  0009-0006-0640-0277  
✉ sevdenurboyraz@gmail.com;  0009-0008-0157-1362.

multiple sensors along a single fiber. These advantages have prompted extensive research into the development of novel fiber optic sensing techniques, including those based on the analysis of speckle patterns.

In recent years, there has been a growing interest in the development of optical fiber bending sensors due to their numerous applications in various fields, including structural health monitoring, robotics, biomechanics, and industrial automation. These sensors utilize the unique properties of optical fibers to detect and measure bending-induced deformations, providing accurate and reliable information about the bending curvature and applied forces. Researchers have used different techniques to measure bending, based on modulation principles such as intensity, wavelength, and frequency modulation [19, 20, 21, 22]. In addition, they have proposed fiber bending devices using different types of fibers such as single mode fiber, multimode fiber, multicore fiber, fiber Bragg grating or long period fiber grating. In a significant part of these studies, bending measurement is performed by calculating the intensity or optical power loss at a standard single or multi-mode fiber output [23, 24]. In others, measurement is performed by analyzing changes in the wavelength spectrum resulting from modal interference. Although bending sensors based on wavelength or frequency modulation are more sensitive to curvature than those based on intensity modulation, the optical spectrum analyzer used in such systems causes a high cost and compromises the compact structure of the sensor.

In this paper, we propose a new fiber bending sensor based on speckle pattern imaging. The use of speckle pattern imaging as a detection mechanism in fiber optic bending sensors offers many advantages. First, it enables non-contact sensing and the quantity to be measured with the sensor eliminates the need for physical contact. Secondly, the speckle pattern imaging technique offers high spatial resolution. Moreover, it will be sufficient to use only one camera instead of using an expensive optical spectrum analyzer.

We organized the paper in four chapter. In the second chapter, the propagation of modes in optical fibers and the formation of the speckle pattern image are presented with a theoretical perspective. In the third chapter, the design and implementation of the optical fiber bending sensor based on speckle pattern imaging is demonstrated by simulated studies. The simulation results and bending sensitivity of the sensor are discussed in the fourth section.

## 2. THE FORMATION OF SPECKLE PATTERN IMAGE IN OPTICAL FIBERS

Optical fibers used in the modern communication industry are the most critical passive optical components that carry the signal in the communication infrastructure. Single-mode fibers are generally preferred for applications that require high data rates or long distances. They are used for very long-distance connections such as

telecommunications, cable TV, applications requiring long-distance data transmission, and even transatlantic undersea link. On the other hand, multimode fibers are often used in many applications such as short-distance local area networks, decorative lighting as a light source, and transporting light from a high-power optical source to the place where it is needed. Fig. 1 (a) and (b) show the distribution of light in the cross-sectional area of single mode and multimode fiber, respectively. As seen in the figure, while in a single-mode fiber the light creates a spot appearance similar to a Gaussian distribution, in a multimode fiber this appearance turns into a complex pattern consisting of many small spots. In fact, this is a result of this fiber carrying multiple light rays or modes simultaneously.

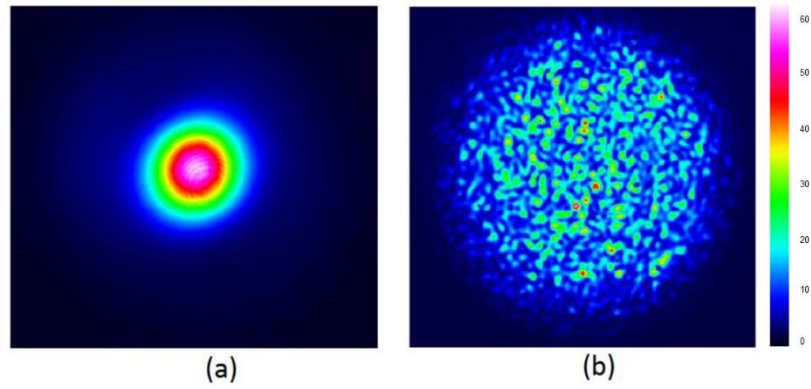


FIGURE 1. Intensity distribution of light at the end of the fiber, (a) Single mode fiber (b) Multimode fiber.

Optical fibers have a radial symmetry. In a multimode step index fiber, hundreds or even thousands of optical modes can be propagated. Since the index profiles exhibit only a small index contrast, they can be assumed to be weakly guiding. In this case, the calculation of fiber modes can be greatly simplified and we can calculate linearly polarized LP modes. The general wave equation in cylindrical coordinates can be written as follows [25].

$$\frac{\partial^2 E}{\partial r^2} + \frac{1}{r} \frac{\partial E}{\partial r} + \frac{1}{r^2} \frac{\partial^2 E}{\partial \varphi^2} + (k^2 - \beta^2)E = 0 \quad (1)$$

Here  $E(r, \varphi)$  is the complex electric field,  $k = 2\pi/\lambda$  is the wave number,  $\beta$  is the phase constant (imaginary part of the propagation constant) and  $\lambda$  is the vacuum wavelength. If a laser beam with an intensity distribution similar to Gaussian is

applied to the input of a multimode fiber ( $z=0$ ), the explanation of the complex electric field is as follows:

$$E_{in}(r, \phi) = \frac{1}{\rho} \exp\left[-\frac{r^2}{\rho^2}\right] \exp(-ik\theta r \cos\phi) \quad (2)$$

Here  $\rho$  is the Gaussian beam radius at the waist and  $\theta$  is the angle of the incident Gaussian beam relative to the optical fiber axis. Any guided field distribution in the multimode fiber can be considered as a superposition of all guided modes. So that the light field traveling along the fiber can be written as the sum of the lights coupled to linear polarization modes:

$$E_{in}(r, \phi) = \sum_m \sum_n \alpha_{mn} LP_{mn}(r, \phi) \quad (3)$$

Here  $\alpha_{mn}$  is the mode field amplitude,  $LP_{mn}$  is any linearly polarized mode and the pair m-n represent to indices of the  $LP_{mn}$  mode. Any  $LP_{mn}$  field is calculated with two separate functions, inside and outside the area surrounded by the fiber core radius. At the entrance of the fiber corresponding to the  $z=0$  position, these fields can be calculated with the following equations using Bessel functions

$$LP_{mn}(r, \phi) = \begin{cases} A \frac{1}{J_m(U)} J_m\left(\frac{U_r}{a}\right) \cos(m\phi) & r \leq a \\ A \frac{1}{K_m(W)} K_m\left(\frac{W_r}{a}\right) \cos(m\phi) & r \geq a \end{cases} \quad (4)$$

Here  $A$  is a constant,  $a$  is the core radius,  $J_m$  and  $K_m$  are the Bessel and modified Bessel functions,  $U$  and  $W$  are normalized transfer propagation constants for  $r \leq a$  and  $r \geq a$ , respectively. The relationships among  $U$ ,  $W$  and normalized frequency  $V$  are defined by the equations given below.

$$\begin{aligned} U &= a(k^2 n_1^2 - \beta^2)^{\frac{1}{2}} \\ W &= a(\beta^2 - k^2 n_2^2)^{\frac{1}{2}} \\ V^2 &= U^2 + W^2 = a^2 k^2 (n_1^2 - n_2^2) \end{aligned} \quad (5)$$

By solving the above equations numerically, we can obtain the intensity or power distribution of each mode. The speckle patterns in optical fibers result from the interference of optical modes that have traveled different paths within the fiber, leading to a random pattern of bright and dark spots as seen in Fig. 1(b). By summing

all the modes guided by the multimode fiber, we can easily calculate numerically the speckle pattern at the incident end of the fiber. The figure 2 shows the patterns of several lower order modes, including the fundamental mode, and the speckle pattern formed by these modes.

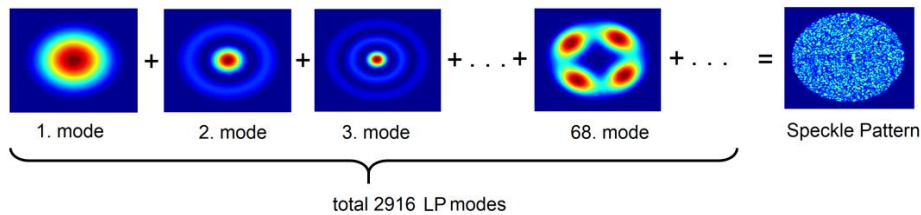


FIGURE 2. Formation of speckle pattern from fiber LP modes.

### 3. DESIGN AND IMPLEMENTATION

The simplest way to measure the response of an optical fiber to bending is to rotate it at diameter  $R$ . Optical fibers can be rotated in loop form up to a few cm in diameter without breaking or permanent deformation, as seen in Figure 3. As the loop radius decreases, a portion of the light escapes from the core to the cladding and is lost, resulting in a decrease in optical power at the fiber output. Although sensors based on the measurement of optical power change at the fiber output are quite practical compared to those using the interferometer principle, they are sensitive to noise and have low resolution due to power oscillations of the light source. Therefore we focused on the change of speckle pattern, not the change of optical power at the fiber output.

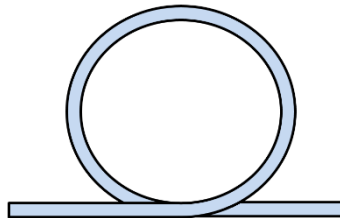


FIGURE 3. Design of fiber bending sensor for a full loop.

In this study, the optical fiber was rotated in the form of a ring with diameter  $R$  and the speckle pattern changes at the fiber end were obtained by using beam propagation method (BPM). This method is very useful for calculating the propagation of light in slowly changing waveguides. However, since the optical field

is calculated with the forward propagation algorithm, the backpropagation of light in waveguides such as ring fiber cannot be calculated. Therefore, we divided a fiber ring into equal arcs of 45 degrees as in Figure 4(a) and we obtained the speckle patterns by simulating in the forward propagation direction of the sensor, which has a form similar to an S-bend. In the simulation studies, we preferred to use a multi-mode fiber with core and cladding diameters of 105/125  $\mu\text{m}$  and  $\text{NA} = 0.22$ . The number of linear modes guided in this fiber is approximately 6000. To obtain the speckle pattern image, simulations must be performed in 3D propagation space. We have experienced that the simulation time for any value of the loop diameter took longer than a week on an ordinary computer. Therefore, by forming the fiber sensor with 6 equal arcs of 10 degrees as shown in Figure 4(b), we reduced the 3D propagation area and completed all simulation studies in a limited time.

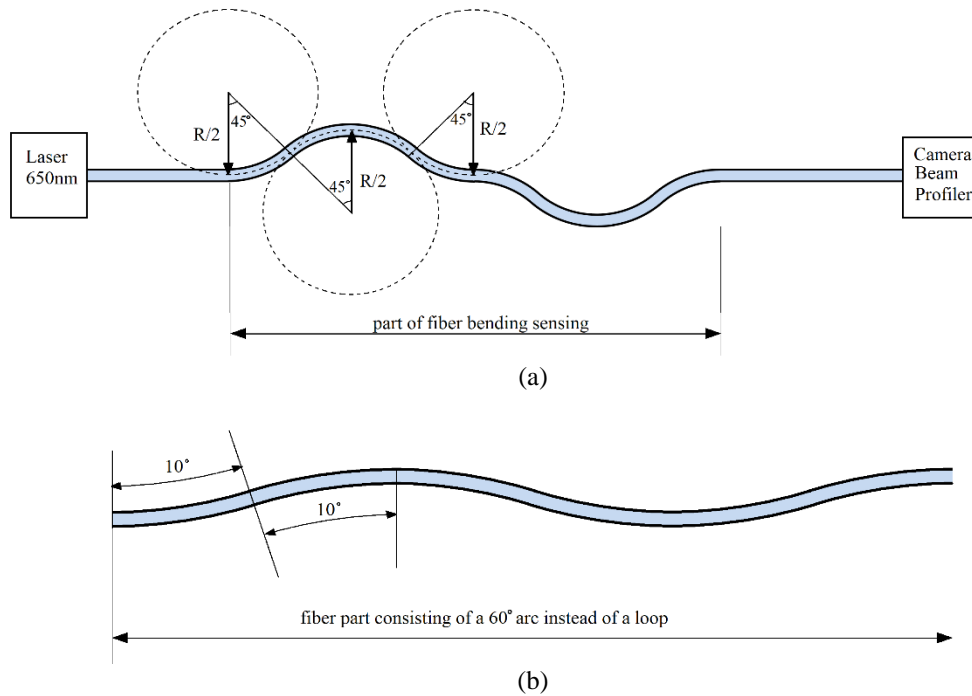


FIGURE 4. Design of S-bending like fiber sensor for simulation implementation in forward propagation a) sensor part whose bending consists of 8 arcs of 45 degrees b) sensor part whose bending consists of 6 arcs of 10 degrees.

In this study, we examined the effect of 60 degree S bending instead of full loop fiber. Of course, bending the fiber 60 degrees around a fixed radius may produce less changes in the speckle pattern than full bending. However, we have shown that bends as small as 60 degrees around different diameters can be detected by fiber speckle pattern imaging. We simulated bends created around different diameters in the range of 1-6 cm and recorded speckle pattern images at the fiber exit. Since approximately 4 cm of straight fiber part is added after bending, the total fiber length increases to 35.306 cm. The complex optical field propagating along the fiber was calculated in a 3D wafer using the BPM method.

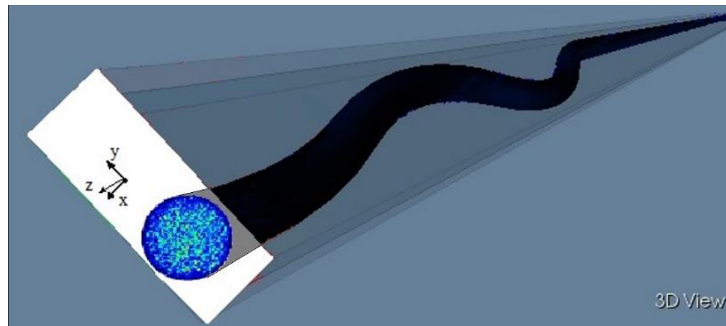
Due to the S structure, the wafer resembles a rectangular prism. Wafer dimensions are adapted to the bending diameter. For a bending diameter of 6 cm, the X-Y cross-sectional area of the wafer is a rectangular plane with dimensions of 1100x126 micrometers as shown in Figure 5(a). Instead of the entire image on the wafer x-y plane, the process was continued with only the part containing the speckle pattern image by cropping and circular masking as shown in Figure 5(b). In addition, image enhancement operations such as background illumination correction and contrast enhancement were performed to achieve uniform illumination and image sharpness. Finally, we quantified how the similarity between consecutive images deteriorated as the bending radius changed by calculating the image correlation.

The cross correlation is a common method to quantify image similarity between two images [26]. As the similarity between images decreases, the correlation coefficient decreases from 1 to 0. Speckle pattern images contain many small-sized speckles. Calculating the average correlation coefficient by segmenting the image instead of the entire image gives better correlation for such images. As seen in Figure 5, the correlation between A and B images whose rows and columns are divided into M and N segments can be calculated as follows [27]:

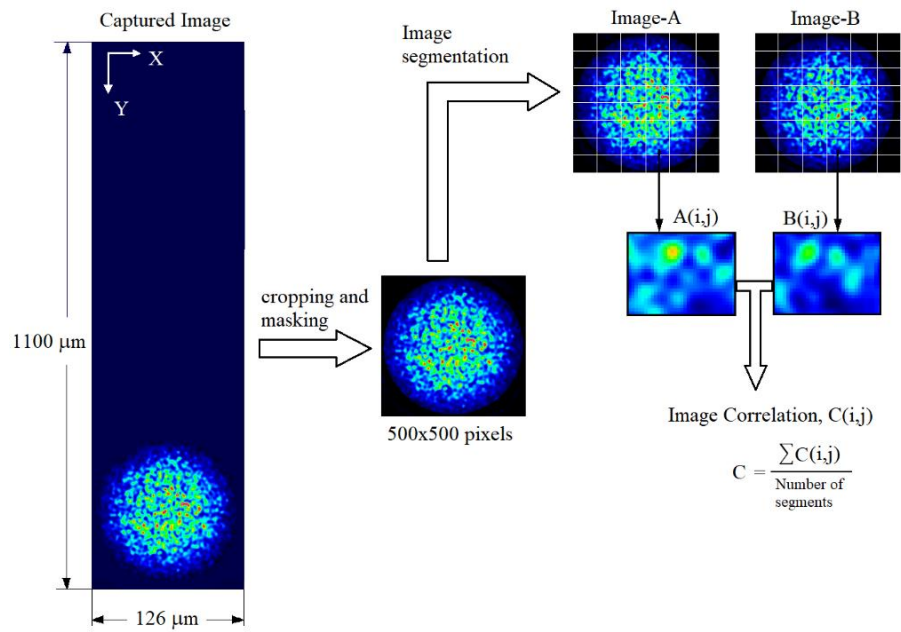
$$C_{i,j} = \frac{\sum(A_{i,j} - \bar{A})(B_{i,j} - \bar{B})}{\sqrt{\sum(A_{i,j} - \bar{A})^2 (B_{i,j} - \bar{B})^2}}$$

$$C = \frac{\sum C_{i,j}}{M \times N}$$

In the next section, we showed that fiber bending can be detected by performing image correlation analysis by applying basic image processing functions.



(a)



(b)

FIGURE 5. (a) 3D view of the wafer, (b) image processing steps.



## 4. RESULT AND DISCUSSION

We started the simulations with fiber bending with a diameter of 6 cm. The first image recorded at the fiber exit with a bending diameter of 6 cm was labeled as the reference image after applying the image processing steps. Then the bending diameter was increased by  $2 \mu\text{m}$  in each new simulation and the last simulation was performed for a bending diameter of 6.0084 cm. Thus, 43 speckle pattern images were obtained. As the bending diameter changes, the sequential speckle pattern images obtained are very similar to each other. It is very difficult to notice the change when followed with the human eye. However, by performing correlation calculation on the segmented image as shown in Figure 5, the change from image to image is clearly visible. Figure 6 shows the variation of the correlation coefficient with a  $2 \mu\text{m}$  diameter increase ( $\Delta R$ ) at each step around the 6 cm bending diameter of the fiber. In the correlation coefficient, the  $C(0)$  value is equal to 1 and this corresponds to the reference image ( $\Delta R=0$ ). After approximately the 40th image, the correlation drops below 0.1. This means that there is very little similarity between the two images.

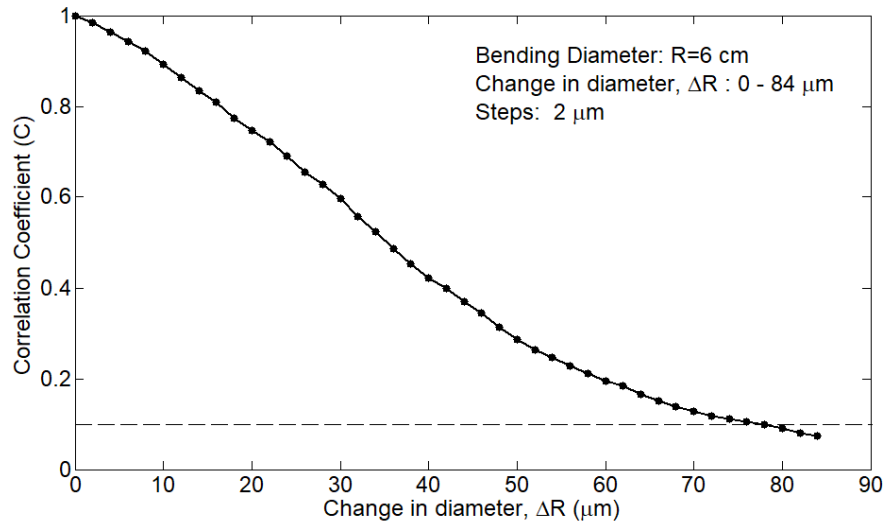


FIGURE 6. The variation of the correlation coefficient versus micrometer changes in bending diameter.

It is clearly seen in Figure 6 that diameter changes at 6 cm fiber bending gradually deteriorate the correlation. This can be explained by the fact that changes in bending

diameter affect the modes propagating along the fiber. However, it is quite remarkable that very small diameter changes such as  $2\ \mu\text{m}$  create measurable changes in the speckle pattern images. This result shows that very small changes in bending diameter can be measured by observing the fiber speckle pattern.

Although the field distributions of all guided modes in a straight fiber can be calculated theoretically, this situation is quite complicated in a bent fiber. Fiber curvature and bend-induced variations in the refractive index both tend to distort the mode field distributions [28]. Therefore, we can predict that as the bending diameter of the fiber decreases, there will be more severe distortions in the speckle pattern image. To verify this, we also tested our sensor for values of the bending diameter in the range of 1-6 cm, respectively. The distortion of correlation in the speckle pattern images was analyzed by changing the bending diameter in small steps of  $2\ \mu\text{m}$ , as in the experiment with bending diameter of 6 cm for this range. Figure 7 shows the variation of the correlation coefficient at different bending diameters.

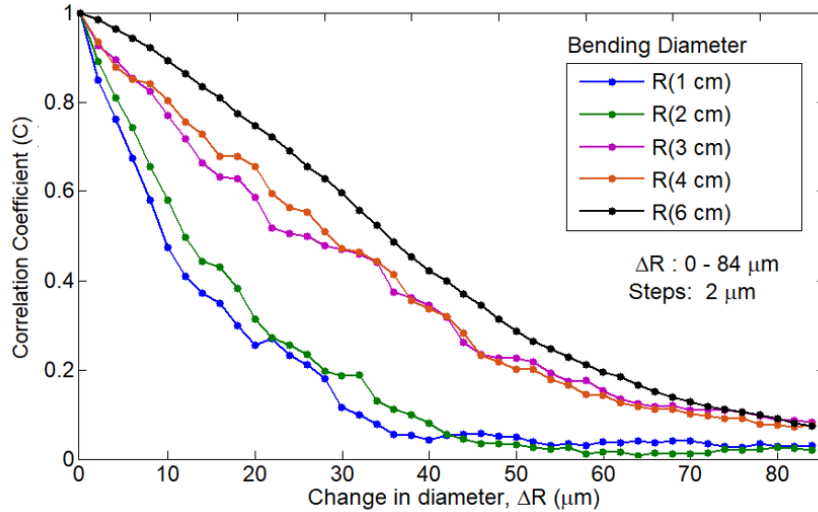


FIGURE 7. The variation of the correlation coefficient for different bending diameter.

It is clearly seen in Figure 7 that as the bending diameter decreases, micro-changes in diameter cause more severe correlation deterioration. The bending sensor with a diameter of 1 cm responds most sensitively to changes in diameter. We can simply define sensor sensitivity by  $|\Delta C/\Delta R|$ . Here,  $\Delta C$  and  $\Delta R$  are the change amounts in correlation and bending diameter, compared to the reference speckle pattern image ( $C = 1$ ). If the diameter of the 6 cm bending sensor changes by  $30\ \mu\text{m}$ , the correlation coefficient drops to 0.5969 and thus its sensitivity is about 0.0013

$\mu\text{m}^{-1}$ . On the other hand, for the bending sensor with a diameter of 1cm, the correlation coefficient drops to 0.1152, so its sensitivity is obtained as  $0.0295 \mu\text{m}^{-1}$ . As a result, the small diameter bending sensor shows more severe correlation changes, which can be used to increase sensor sensitivity.

## 5. CONCLUSIONS

In this paper, a fiber bending sensor is proposed based on the correlation distortion in speckle pattern images. The bending sensor designed using a multimode fiber curved 60 degrees in a form similar to the S structure. The sensitivity of the sensor with different curvature diameters in the range of 1-6 cm was tested. The sensitivity was found to be  $0.0013 \mu\text{m}^{-1}$  at 6 cm bending diameter and  $0.0295 \mu\text{m}^{-1}$  at 1 cm bending diameter. It has been shown that the sensitivity increases as the bending diameter decreases. Our sensor can detect changes in the bending diameter at the  $\mu\text{m}$  level. We have demonstrated detection with only 60 degrees of S-structure bending. If the sensor structure is designed in the form of a loop, detection will be even more sensitive. On the other hand, the detection process is quite different from the methods applied in traditional fiber sensors. The detection principle of the proposed sensor is based on the processing of images with fiber speckle patterns, which offers a significant advantage for characterizing the sensor. Moreover, the sensor can be trained with deep learning algorithms and perform accurate measurements.

**Author Contribution Statements** The authors equally worked on this study. All authors contributed to this study at every stage.

**Declaration of Competing Interests** The authors declare that there is no conflict of interest regarding the publication of manuscript.

**Acknowledgement** This research was supported by the Scientific and Technological Research Council of Turkey (TUBITAK 2209-A Program in 2023).

## REFERENCES

- [1] Gao, H., Hu, H., Zhao, Y., Li, J., Lei, M., Zhang, Y., Highly-sensitive optical fiber temperature sensors based on PDMS/silica hybrid fiber structures, *Sens. Actuators A Phys.*, 284 (2018), 22-27, <https://doi.org/10.1016/j.sna.2018.10.011>.
- [2] Su, H., Zhang, Y., Ma, K., Zhao, Y., Wang, J., High-temperature sensor based on suspended-core microstructured optical fiber, *Opt. Express*, 27 (2019), 20156, <https://doi.org/10.1364/OE.27.020156>.

- [3] Li, M., Gong, Y., Yin, J., Li, W., Shao, Y., Cong, A., Huang, G., Highly-sensitive and wide-range temperature sensor based on polymer-filled micro-cavity in fibre Bragg grating by temperature segmentation, *Optik*, 245 (2021), 167707.
- [4] Sun, X., Zhang, L., Zeng, L., Hu, Y., Duan, J., Micro-bending sensing based on single mode fiber spliced multimode fiber Bragg grating structure, *Opt. Commun.*, 505 (2022), 127513, <https://doi.org/10.1016/j.optcom.2021.127513>.
- [5] Perez-Herrera, R.A., Andre, R.M., Silva, S.F. et al., Simultaneous measurement of strain and temperature based on clover microstructured fiber loop mirror, *Measurement*, 65 (2015), 50-53, <https://doi.org/10.1016/j.measurement.2014.12.052>.
- [6] Bilsel, M., Navruz, I., Tapered optical fiber sensor for discrimination of strain and temperature, *Advances in Electrical and Electronic Eng.*, 18 (2020), 50-56.
- [7] Kissinger, T., Correia, R., Charrett, T. O. H., James, S. W., Tatam, R. P., Fiber segment interferometry for dynamic strain measurements, *J. Light. Technol.*, 34 (2016), 4620-4626, <https://doi.org/10.1109/JLT.2016.2530940>.
- [8] Sazio, P. J. A., Microstructured optical fibers as high-pressure microfluidic reactors, *Science*, 311 (2006), 1583-1586.
- [9] Dong, N., Wang, S., Jiang, L., Jiang, Y., Wang, P., Zhang, L., Pressure and temperature sensor based on graphene diaphragm and fiber Bragg gratings, *IEEE Photonics Technol. Lett.*, 30 (2018), 431-434, <https://doi.org/10.1109/LPT.2017.2786292>.
- [10] Zhang, W., Ni, X., Wang, J., Ai, F., Luo et al., Microstructured optical fiber based distributed sensor for in vivo pressure detection, *J. Lightwave Technol.*, 37 (2019), 1865-1872.
- [11] Kim, H. J., Shin, H. Y., Pyeon, C. H., Kim, S., Lee, B., Fiber-optic humidity sensor system for the monitoring and detection of coolant leakage in nuclear power plants, *Nucl. Eng. Technol.*, 52 (2020), 1689-1696.
- [12] Bian, C., Wang, J., Bai, X., Hu, M., Gang, T., Optical fiber based on humidity sensor with improved sensitivity for monitoring applications, *Opt. Laser Technol.*, 130 (2020), 106342.
- [13] Zhang, J., Shen, X., Qian, M., Xiang, Z., Hu, X., An optical fiber sensor based on polyimide coated fiber Bragg grating for measurement of relative humidity, *Opt. Fiber Technol.*, 61 (2021), 102406, <https://doi.org/10.1016/j.yofte.2020.102406>.
- [14] Huang, X. Lai, M., Zhao, Z., Yang, Y. et al., Fiber optic evanescent wave humidity sensor based on SiO<sub>2</sub>/TiO<sub>2</sub> bilayer films, *Appl. Opt.*, 60 (2021), 2158-2165.
- [15] Wang, T., Yasukochi, W., Korposh, S., James, S. W., Tatam, R. P., Lee, S.-W., A long period grating optical fiber sensor with nano-assembled porphyrin layers for detecting ammonia gas, *Sens. Actuators B*, 228 (2016), 573-580.
- [16] Yu, C.-B., Wu, Y., Li, C., Wu, F., Zhou, J.-H., Gong, Y., Rao, Y.-J., Chen, Y.-F., Highly sensitive and selective fiber-optic Fabry-Perot volatile organic compounds sensor based on a PMMA film, *Opt. Mater. Express*, 7 (6) (2017), 2111-2116.
- [17] Sultangazin, A., Kusmangaliyev, J., Aitkulov, A., Akilbekova, D., Olivero, M., Tosi, D., Design of a smartphone plastic optical fiber chemical sensor for hydrogen sulfide detection, *IEEE Sens. J.*, 17 (21) (2017), 6935-6940.

- [18] Hosok, A., Nishiyama M., Kumekawa N., Watanabe, K. Et al., Hetero-core structured fiber optic chemical sensor based on surface plasmon resonance using Au/lipid films, *Opt. Commun.*, 524 (2022), 128751, <https://doi.org/10.1016/j.optcom.2022.128751>.
- [19] Wu, Y., Pei, L., Jin, W., Youchao, J., Yang, Y., et al., Highly sensitive curvature sensor based on asymmetrical twin core fiber and multimode fiber, *Opt. Laser Technol.*, 92 (2017), 74-79, <https://doi.org/10.1016/j.optlastec.2017.01.007>.
- [20] Gong, Y., Zhao, T., Rao, Y-J., Wu, Y., All-fiber curvature sensor based on multimode interference, *IEEE Photonics Technol. Lett.*, 23 (2011), 679-681.
- [21] Li, Y-P., Zhang, W-G., Wang, S., Chen, J. et al., Bending vector sensor based on a pair of opposite tilted long-period fiber gratings, *IEEE Photonics Technol. Lett.*, 29 (2017), 224-227, <https://doi.org/10.1109/LPT.2016.2636446>.
- [22] Chen, Y., Yu, Z., Chen, H., Tao, C., et al., Experimental study on temperature-insensitive curvature sensor based on reflective all-fiber structure, *Infrared Phys. Techn.*, 137 (2024), 105146, <https://doi.org/10.1016/j.infrared.2024.105146>.
- [23] Anderson, D. Z., Bolshtyansky, M. A., and Zel'dovich, B. Y., Stabilization of the speckle pattern of a multimode fiber undergoing bending, *Opt. Lett.*, 21 (11) (1996), 785-787.
- [24] Asawa, C. K., Taylor, H. F., Propagation of light trapped within a set of lowest-order modes of graded-index multimode fiber undergoing bending, *Appl. Opt.*, 39 (2000), 2029-2037.
- [25] Keiser, G., *Optical Fiber Communication*, Mc Graw Hills, Third Edition, Singapore, 2000.
- [26] Schreier, H., Orteu, J-J., Sutton, M. A., *Image correlation for shape, motion and deformation measurements*, Springer, 2009.
- [27] Ari, F., Serbetci, H., Navruz, I., Tapered fiber optic refractive index sensor using speckle pattern imaging, *Opt. Fiber Technol.*, 79 (2023), 103366.
- [28] Schermer, R. T., Mode scalability in bent optical fibers, *Optics Express*, 15 (24) (2007), 15674-15701, <https://doi.org/10.1364/OE.15.015674>.

## SENDING PICTURES OVER RADIO SYSTEMS OF THE TRAIL CAM IN BORDER SECURITY AND DIRECTING UAVS TO THE RIGHT AREAS

Vedat YILMAZ<sup>1</sup>


<sup>1</sup>Institute of Forensics Sciences, Department of Criminalistics, JSGA, Ankara, TÜRKİYE

**ABSTRACT.** In this study, a method is proposed for the trail cams to send data via narrow band communication systems in border security and counter-terrorism areas and to direct drones to the right areas. The success of UAVs lies in scanning the correct areas for observation or detection. UAVs should be fed with data to observe the correct regions, and the probability of detecting border security or terrorist elements should be increased. Instantaneous detection is performed by trail cam, which generally operate dependent on GSM. However, these devices cannot provide real-time data in border areas with low population density and no GSM service, particularly in counter-terrorism operations. In this study, the dependence of trail cam devices on GSM was eliminated, and data transfer over the radio system was established to enable real-time data flow in a wide field. After the trail cam device makes a detection, the data is sent via the APCO-25 JEMUS radio system with a capacity of 9.6 KB. The resolution of the detection image is reduced, allowing it to be displayed on a remote-control computer in less than one minute. As a result of the study, when an intelligent trail cam with object recognition capability is developed, the device can assess what the image might be in real-time. Obtaining real-time detection data from trail cams in border areas and counter-terrorism zones without GSM infrastructure can expedite the direction of UAVs to the correct regions for intervention by military units. Additionally, confirming that trail cam detects via narrowband communication systems in locations where units are temporarily stationed and without alpine terrain minimizes the surveillance vulnerability of UAVs unable to perform imaging due to adverse weather conditions. This also establishes a warning system against potential attacks by terrorist elements.

### 1. INTRODUCTION

Trail cams are actively used worldwide, especially to observe natural wildlife [1]. In our country, these devices are not only utilized for observing natural life but are also actively employed by the armed forces and law enforcement agencies for detection purposes in the context of border security and counter-terrorism [2]. Trail cams are

*Keywords.* Border security, trail cam, detection, reconnaissance, surveillance.

✉ vedat.yilmaz@jsga.edu.tr-Corresponding Author;  0000-0002-3112-9371.

equipped with sensors to detect motion or heat, and they feature specialized cameras [3]. These cameras are automatically triggered when they detect movement or a specific temperature difference [1, 2, 3]. Upon triggering, they capture photographs or videos. However, these devices have limitations such as weather and environmental factors, energy constraints, false triggers, and data processing challenges [4]. Furthermore, the current limitation of these devices lies in their reliance on GSM-based transmission for real-time visual communication. This reliance restricts their usage within the GSM coverage area or confines them to storing images in their own memory outside the coverage area [5]. The devices storing detection data on themselves leads to delayed learning by the intervention unit, hindering prompt action [1, 3, 4, 5, 6]. Additionally, when terrorists detect the devices or realize they are being tracked, they may booby-trap the devices with improvised explosive devices [7]. This poses a risk to personnel approaching the devices to retrieve data, resulting in injuries or fatalities [8].

To address these issues, new satellite-based camera trap devices are being developed to ensure continuous data transmission [7, 8]. Smart trail cams with satellite-based data transmission capability use on-device image processing algorithms for detection and provide wide coverage [7, 8, 9, 10,11]. However, the use of satellite transmission increases the cost of these devices, limiting their deployment to critical areas with a constrained number of devices [8]. Instant data transmission by trail cam is a crucial factor in minimizing casualties while combating terrorism, ensuring operational success, and responding promptly to incidents [7]. In areas without GSM infrastructure, narrowband communication systems are commonly used for communication [12, 13, 14]. Ensuring real-time visual transmission via these narrowband communication systems can address the challenges and minimize personnel vulnerabilities [13].

Narrowband communication systems generally refer to radio communication, providing communication within a limited bandwidth [14]. The coverage area of radio-based communication systems is significantly larger compared to GSM, especially when considering the eastern and southeastern borders of the Republic of Turkey and counter-terrorism operation zones, where the population density is low, leading to limited GSM infrastructure and coverage [15]. The existing coverage of narrowband (radio) systems in these areas is quite high, and it can be expanded in desired areas using mobile repeater centers [16]. Additionally, digital radio systems not only facilitate voice communication but also offer limited capabilities for location transmission and data transfer [15]. For instance, radios operating according to the Apco-25 standards provide a bandwidth of 9.6 kilobytes [15, 17].

In today's world, unmanned aerial vehicles (UAVs) used in border security operations are continually evolving to maximize the advantages of technology [18]. UAVs play an effective role in border areas, particularly due to the advantages they

offer in reconnaissance and surveillance [19, 20]. The success of UAVs lies in scanning the correct areas for observation and detection [19]. UAVs need to be fed with data to observe the correct regions, and the probability of detecting border security or terrorist elements should be increased [21, 22].

When examining the advantages of UAVs in border security and counter-terrorism operations, reconnaissance and surveillance functions come to the forefront [23]. In situations where traditional surveillance methods fall short, UAVs can quickly scan a large area using high-resolution cameras and sensors to detect potential threats [24]. UAVs can maneuver rapidly and flexibly in border areas. Operating at a lower cost compared to traditional aircraft, UAVs can perform longer and more regular missions in border security operations, enabling more effective monitoring of border regions. UAVs can detect and analyze potential dangers in border areas [25]. Equipped with technological features such as thermal cameras, night vision systems, and radars, UAVs can identify smuggling, terrorist activities, and other potential threats in border areas, assisting in the implementation of preventive measures [26, 27]. The use of UAVs in border security operations can enhance personnel safety [24]. UAVs operating in dangerous or hard-to-reach areas do not jeopardize human safety [22, 23, 24]. Additionally, with a lower operational cost compared to traditional aircraft, UAVs allow for a more efficient utilization of border security budgets [24, 28]. To further enhance the effectiveness of UAVs by flying them in the correct areas, taking into consideration remote points in borders and settlements with natural vegetation, trail cam devices can serve as the "eyes" on the ground for UAVs. With advancements in imaging technologies, trail cam devices can play a critical role in providing data in border areas and directing UAVs to the right locations.

In this way, national border protection and illegal crossings will be prevented, the effectiveness of the fight against terrorists will increase, and effective use will be ensured by bringing together intelligent systems and controls in terms of national security. The aim of this study is to propose a method for the use of trail cams in wide-field applications in border security and counter-terrorism regions. This involves the real-time transmission of data through underutilized narrowband systems and the deployment of UAVs in the correct areas for control and detection.

## 2. LITERATURE REVIEW AND CURRENT SITUATION

To transmit image data over a radio system, it is necessary to first convert this image into digital data [29]. This process is typically carried out using a video compression algorithm. Commonly used algorithms include H.264, H.265 (HEVC), and MPEG-4 [30]. Radio systems are generally designed for voice communication [12-15, 31]. Therefore, it may be necessary to initially convert image data into audio data [31-



33]. This process involves carrying the image data over a voice communication channel and then converting it back into an image on the receiving end [14]. This often includes transforming the image data into a format compatible with voice communication protocols [31]. Once the image data has been transported over the voice communication channel, this audio data is sent to another center via a radio frequency [33]. This typically occurs in the form of a radio data packet. On the receiving end, the image data is received as audio data [14]. Subsequently, this audio data is processed to be transformed back into the original image data. This step involves using video compression algorithms utilized to convert audio data back into an image [34]. This process is quite complex, especially concerning the specialized protocols and algorithms used in radio communication systems. Additionally, issues of privacy and security are crucial in such communication, hence specialized encryption methods ensuring secure communication are often employed.

### 2.1. Steps for Data Transmission Over Radio.

#### a. Matrix Transformation of the Image;

Let's consider a matrix representing the image as  $I$ . Each element  $I_{i,j}$  represents a pixel of the image. This matrix typically has three channels according to a color space (e.g., RGB).

$I$ : Original image matrix (dimensions  $m \times n$ )

$I_{i,j}$ : Element of the original image matrix (dimensions  $m \times n$ )

#### b. Video Compression;

If a video compression algorithm is used, it usually employs a transformation matrix  $C$ . This transformation matrix compresses the original image matrix into a compressed format.

$C$ : Compression matrix (usually a transformation matrix)

$I'$ : Compressed image matrix

$$I' = C \times I \quad (1)$$

#### c. Conversion to Audio Data;

The compressed image matrix is then transmitted over a voice communication channel. This communication typically uses a protocol designed for audio data.

$f$ : Function that converts the compressed image matrix to audio data

$S$ : Audio data

$$S = f(I') \quad (2)$$

d. Transmission via Radio: Audio data is transmitted to another center over a specific frequency by the radio system. This process typically occurs using a radio communication protocol.

$T$ : Transmission matrix (transmission over the radio)

$$T(S) \quad (3)$$

e. Processing at the Receiver Radio: After receiving the audio data at the receiver side, this audio data is first transformed back into the original image matrix. This process can be mathematically expressed as follows:

g: Function that transforms audio data to the original image matrix

"I": Original image matrix obtained at the receiver side

$$I''=g(T(S)) \quad (4)$$

Communication devices conforming to the APCO-25 standard can be connected to tablets or computers via connectors for the purpose of data transmission when desired [32, 33]. Data transfer is facilitated through data relays located in repeater centers. Additionally, the location data of these radio devices can be viewed from the central control software through the GPS module [32]. The data transmission duration from the radio device is dependent on the size of the transmitted data. Therefore, reducing the size of the sent image will decrease the transmission time. The process of reducing image size is typically achieved using data compression algorithms. The mathematical formulation of these algorithms varies depending on the specific details of the algorithms employed and the compression method used. The JPEG compression algorithm, for instance, typically compresses the original image through a process known as Discrete Cosine Transform (DCT) and quantization. These processes can be mathematically expressed as follows.

2.2. **Discrete Cosine Transform (DCT) Process  $F(u,v)$ .** DCT coefficients in the frequency domain of the original image  $f(x,y)$ : Pixel values in the time domain of the original image M and N: Dimensions of the image,

$$F(u, v) = C(u)C(v) \sum_{y=0}^{N-1} x \sum_{x=0}^{M-1} f(x, y) \cos \left[ \frac{(2x+1)u\pi}{2M} \right] \cos \left[ \frac{(2y+1)v\pi}{2N} \right] \quad (5)$$

c. Quantization Process:

Q(u,v): Quantization table

Fq(u,v): Quantized DCT coefficients

$$Fq(u, v) = \frac{F(u,v)}{Q(u,v)} \quad (6)$$

These steps involve obtaining the Discrete Cosine Transform (DCT) coefficients of the original image in the frequency domain, followed by the quantization process to represent these coefficients with lower precision [34]. The JPEG algorithm compresses these quantized coefficients further through additional steps such as zigzag scanning and Huffman coding [34, 35]. This formulation only exemplifies the JPEG compression algorithm. For other compression algorithms, different

mathematical expressions may arise depending on the methods and techniques employed [34, 35].

### 3. MATERIAL AND METHOD

In the conducted study, two radio devices, one receiver and one transmitter, compliant with the Apco-25 standard, one trail camera, connection cables, and one tablet were utilized. The transmitter and the trail camera were interconnected and placed in a forested area. The location data of the trail camera was obtained through the radio device. The acquired location of the trail camera was marked on a map within the central control software. When the trail camera captured an image through triggering, it stored the image locally and transmitted information about the detection via a short message to the receiver radio device. The central control software installed on the tablet connected to the receiver radio device issued a notification of the detection, including the trail camera's identification number and the date and time of the event. The user personnel could view the notification and, at their discretion, initiate the process of locking the reduced-resolution image data to be transmitted by the transmitter radio device. This allowed the image data to be visible on the control computer.

### 4. THE EXPERIMENTAL SECTION AND DISCUSSION

Unmanned Aerial Vehicles (UAVs), also known as drones, provide the capability to rapidly diagnose potential threats in border regions and operational zones [19, 23, 24]. To ensure effective visual monitoring of violations and terrorist elements and to enable swift intervention within the legal framework, UAVs must operate in the correct areas. Instant visual transmission by trail cameras in the field will offer significant advantages to armed forces and law enforcement in maintaining dominance and detecting irregular migration movements by directing UAVs to the right areas. Trail camera, equipped with high-resolution cameras and sensitive sensors, have the capacity to obtain detailed images in border regions. These devices can operate effectively in both day and night conditions, providing valuable data in various weather conditions.

While these devices are capable of obtaining high-resolution images or videos, for the military and law enforcement, the information that there is a violation and detection in the captured image is more crucial in the initial stage for prompt intervention. This is because responding to a violation or a terrorist element requires a rapid and efficient process that operates against time. Therefore, promptly reporting the detection to the control center is of vital importance.

The ability of trail cameras to transmit data via radio is often encountered as a technology used for remote control and data collection. While these devices are typically used for observation and monitoring in nature, their remote control and real-time data transfer are facilitated through radio transmission.

For camera traps to perform data transmission via radio systems, the device must establish a connection with a control unit in the field or a main station that enables remote access for the user. Additionally, data security during radio transmission is another aspect that needs to be considered for such devices.

During radio transmission, the security of data is of paramount importance. End-to-end encryption is a critical element for data security. By employing encryption during data transmission, protection against threats such as unauthorized access and data manipulation is ensured.

The radio transmission of trail cameras provides users with the ability for remote control and monitoring in areas without GSM infrastructure. This allows users located remotely from the device's position to view live footage, adjust camera settings, and control the device remotely.

The use of trail cameras must comply with legal regulations and ethical standards. Special attention should be paid to matters such as personal privacy, protection of private spaces, and conservation of natural habitats.

**4.1. Advantages of Radio Transmission of Trail Cameras.** Trail Cameras offer various advantages in security applications when utilizing radio transmission.

a. **Area Monitoring and Motion Detection:** Trail Cameras can detect potential threats by continuously monitoring a specific area through their motion detection capabilities. This feature is crucial, especially in areas such as border security, counter-terrorism, facility security, and wildlife conservation. Combined with motion detection, security personnel can be instantly informed, enabling swift intervention.

b. **Image and Video Recording:** Trail Cameras typically feature integrated cameras capable of high-resolution image and video recording. These recordings can be used to document, substantiate, and analyze security incidents. Additionally, these images can be used to assess the effectiveness of measures taken during an event.

c. **Resilience to Weather and Environmental Conditions:** Camera traps are generally weather-resistant and can be adapted to various environments with suitable casings. This adaptability allows effective use in applications such as counter-terrorism and border control.

d. **Covert and Remote Positioning:** Trail Cameras used in security applications are often concealed or camouflaged, allowing them to go unnoticed while detecting and recording potential threats. The remote controllability of trail cameras enables security personnel to manage devices in areas with difficult access or potential risks.

e. **Rapid Intervention and Interactive Control:** The real-time information provided by camera traps enables quick intervention and interactive control. Combined with motion detection, security personnel can promptly dispatch airborne elements or intervention personnel to the area. If necessary, devices can be remotely controlled, providing security teams with the ability to respond quickly and effectively.

f. **Energy Efficiency:** One of the significant advantages of camera traps transmitting notifications via radio is the use of low-power radio frequency (RF) transmission. This contributes to energy efficiency, allowing for prolonged device usage. RF transmission typically occurs in dedicated frequency bands, reducing interference with other wireless devices. Additionally, the ability of smart camera traps to wake up and perform object or person recognition upon motion detection minimizes unnecessary data transmission and energy consumption. Designing suitable energy sources such as solar panels or battery packs for these devices can further enhance their long-term field use. Energy-efficient sensors and electronic components also contribute to energy savings.

**4.2. Disadvantages of Radio Transmission of Trail Cameras.** Trail Cameras have the following disadvantages.

a. **Cost Factors and Limitations of Widespread Use** The cost of high-quality camera traps and radio transmission systems may limit the widespread adoption of this technology. Establishing a network spread across large areas for observing wildlife, conducting research, or for security purposes can impose a substantial financial burden. However, considering the geography of Turkey, military units in border areas, especially in operations related to internal security and counter-terrorism, have existing radio communication infrastructure along the border lines.

b. **Sensor Sensitivity and Environmental Challenges** Accurate operation of camera traps requires sensitive sensors. However, environmental conditions, weather, and natural factors can affect sensor performance. Factors such as dense vegetation, precipitation, and temperature fluctuations can lead to sensors providing incorrect positive or negative responses, reducing the reliability of the obtained data.

c. **Connectivity Issues** Radio transmission can be problematic in remote areas or areas with dense vegetation. This can result in devices being unable to reliably transmit data to the control center, preventing users from effectively managing the device. Particularly in remote and isolated areas, issues like blockage or attenuation of radio signals can challenge the reliability of these systems.

d. **Energy Management and Battery Life** Considering that camera traps are often used in natural environments, energy sources become a critical factor. While energy sources such as solar panels or battery packs are used, these sources have limited capacities. Especially in situations requiring prolonged observation and data collection, special attention must be given to energy management and battery life.

Following the detection process, a notification of the detection was sent via short message through the radio data transmission channel. The receiver radio device, connected to the tablet, displayed an alert of the detection using the control software. After receiving the detection alert, a request for image data was sent to the transmitter radio device via short message. The process of compressing and reducing the resolution of the image data was initiated on the camera trap device, and the transmission of the compressed image data to the receiver radio device was started by the transmitter radio device. The entire process ensured that the reduced-resolution image data, approximately 400 to 500 megabytes in size, was visible in the control software within a time frame of approximately 45-60 seconds. The process is illustrated in Figure 1.

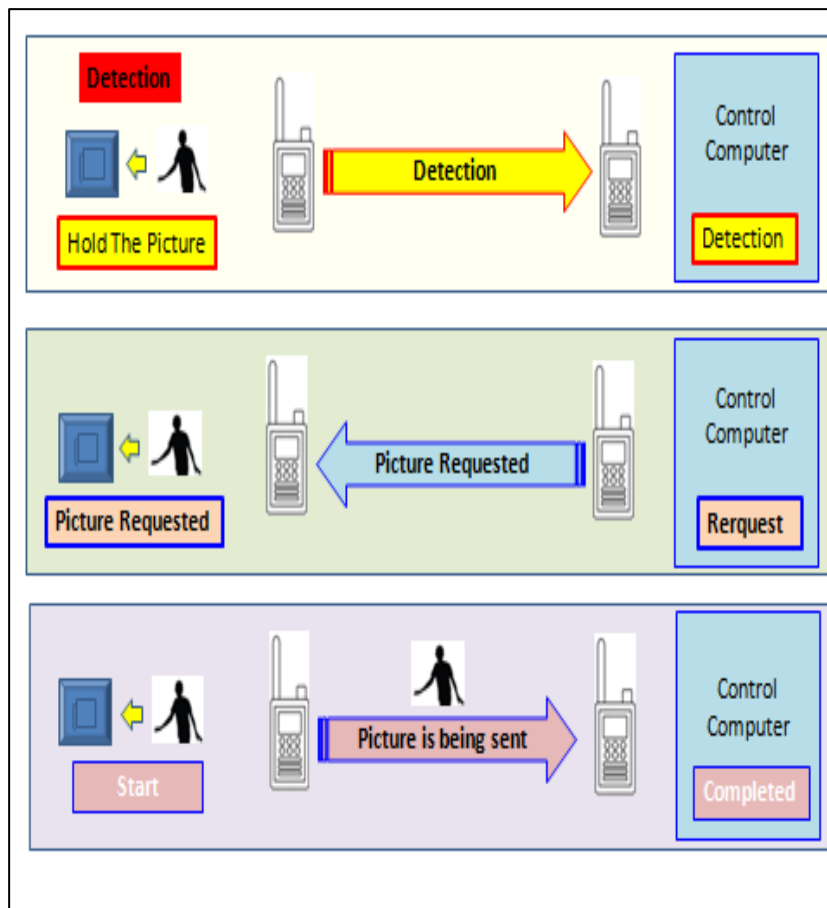


FIGURE 1. Post-detection flowchart.

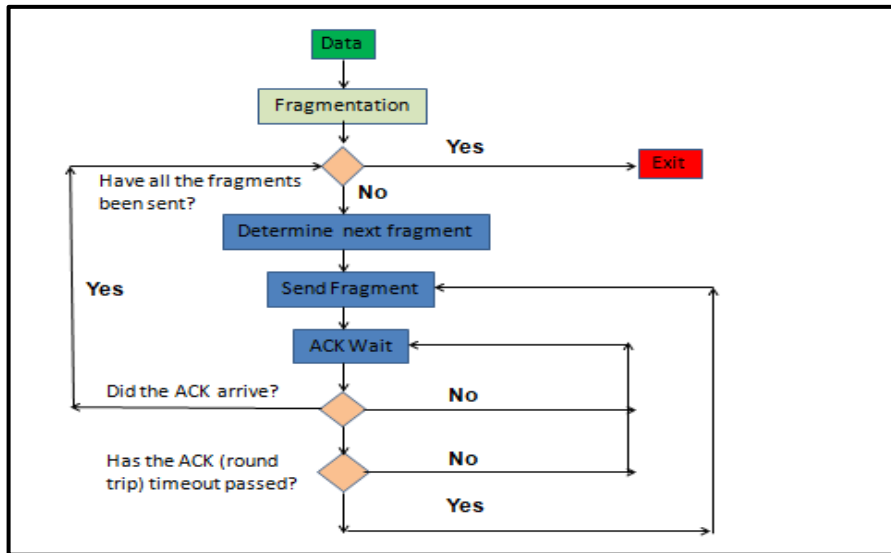


FIGURE 2. Data transmitting radio (ACK= Acknowledgment).

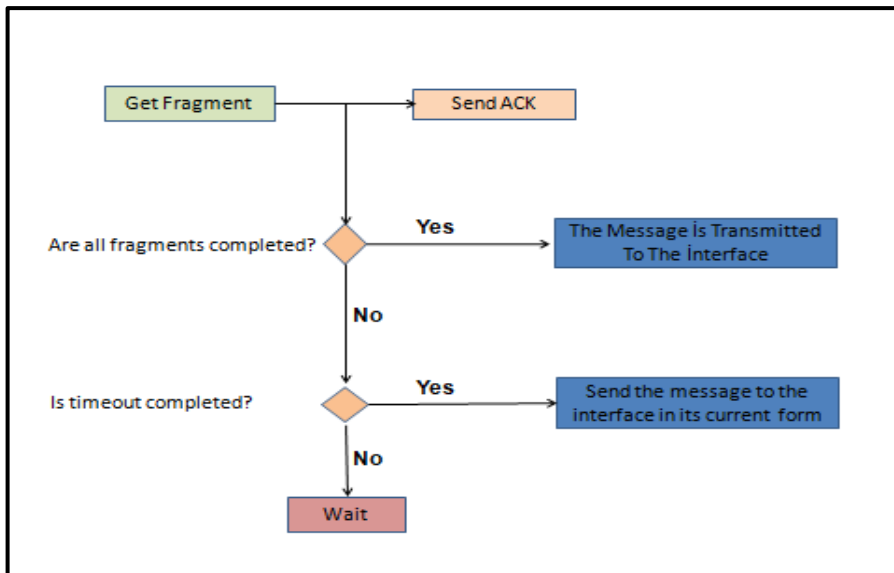


FIGURE 3. Data receiving radio (ACK = Acknowledgment).

Conducted in the experimental study, the process carried out by the camera trap device following the detection is depicted in Figure 2. The operation performed by the receiver radio device is illustrated in Figure 3.

The operation of trail cam systems, particularly in border regions, via radio systems for data transmission, is crucial for ensuring the perimeter security of stationary units. Attacks on these stationary elements typically occur in conditions where visibility is limited, and unmanned aerial vehicle (UAV) elements cannot conduct surveillance due to adverse weather conditions, such as cloud cover. In this context, it can be utilized as a security measure for the unit and as an early warning system with advanced surveillance capabilities. Similarly, in densely forested areas, UAV surveillance can pose challenges.

The image obtained by the camera trap device connected to the radio system in a forested area is depicted in Figure 4.



FIGURE 4. The image captured by the camera trap is transmitted through the radio device.

After the conducted study, the radio system with a data transmission capacity of 9.6 kbytes was utilized to transmit image data. By reducing the resolution of the image, it was ensured that the image data, ranging between 350-500 kbytes, could be captured from the desired camera trap device in less than approximately 1 minute when needed.

## 5. CONCLUSION

The study results indicate that with the development of an intelligent trail camera equipped with on-site object recognition capabilities, the device responsible for recognition can initially assess what the image might contain in real-time. This allows the radio device to provide preliminary information to the user regarding the detection through the radio device. Based on the user's preliminary assessment, the detection image can be viewed within approximately one minute. This aspect has been demonstrated in the established test setup. Consequently, limitations such as GSM coverage restrictions or the disadvantage of holding images that cannot be obtained instantly on the trail camera can be overcome. This enables the detection of



irregular migration movements, border crossings, and the identification of terrorist elements in areas where trail cameras cannot be used due to their limitations. As a result, UAVs can conduct reconnaissance and surveillance in the right areas. Additionally, the technology can be utilized for the safety and advanced surveillance of locations where UAVs are temporarily unable to provide surveillance due to adverse weather conditions, offering significant advantages to the Armed Forces and law enforcement.

While the transmission advantages of trail cameras through radio signals are evident, disadvantages such as compliance with legal regulations and concerns about privacy should also be taken into consideration. To effectively and responsibly use this technology, users must exercise caution and adhere to local regulations. The integration of trail cameras with UAVs has the potential to enhance border security operations, making them more efficient, rapid, and secure. This technological integration is seen as a crucial step in increasing national security and monitoring border areas more effectively. Therefore, trail cameras should be used via GSM-based communication in areas with GSM infrastructure and through radio or satellite-based communication in critical and essential areas where GSM infrastructure is not available for real-time visual data transmission and extensive field use.

**Declaration of Competing Interests** The authors declare no conflict of interest.

#### REFERENCES

- [1] Lupp, G., Kantelberg, V., Förster, B., Honert, C., Naumann, J., Markmann, T., and Pauleit, S, Visitor counting and monitoring in forests using camera traps: A case study from Bavaria (Southern Germany), *Land*, 10 (7), (2021), 736, <http://dx.doi.org/10.3390/land10070736>.
- [2] Şimşek, M., Yalçinkaya, F., Uğurlutan, R., Wide area scanning Trap Camera System with multi-cameras and distinctive motion detection sensor, *26th SIU IEEE*, (2018), 1-4, <https://doi.org/10.1109/SIU.2018.8404190>.
- [3] Albers, J. L., Wildhaber, M. L., Green, N. S., Struckhoff, M. A., and Hooper, M. J., Visitor use and activities detected using trail cameras at forest restoration sites, *Ecological Restoration*, 41 (4) (2023), 199-212, <https://doi.org/10.3368/er.41.4.199>.
- [4] Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., and Clune, J., Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning, *Proc. Natl. Acad. Sci.*, 115 (25) (2018), E5716-E5725, <https://doi.org/10.1073/pnas.1719367115>.
- [5] Schneider, S., Taylor, G. W., & Kremer, S. (2018, May). Deep learning object detection methods for ecological camera trap data, *15<sup>th</sup> CRV IEEE*, (2018), 321-328, <http://dx.doi.org/10.1109/CRV.2018.00052>.

- [6] Şimşek, E., Özyer, B., Bayındır, L., Özyer, G. T., Human-animal recognition in camera trap images, *26<sup>th</sup> SIU IEEE*, (2018), 1-4, <http://dx.doi.org/10.1109/SIU.2018.8404700>.
- [7] Roboteye, (2023). Available at: <https://roboteye.ai/kapan-solo/>. [Accessed January 2024].
- [8] Rpboteye, (2023). Available at: [https://www.linkedin.com/posts/roboteyeai\\_join-us-at-the-world-defense-show-in-riyadh-activity-7158686794442539010-5K00](https://www.linkedin.com/posts/roboteyeai_join-us-at-the-world-defense-show-in-riyadh-activity-7158686794442539010-5K00). [Accessed January 2024].
- [9] Yousif, H., Yuan, J., Kays, R., He, Z., Fast human-animal detection from highly cluttered camera-trap images using joint background modeling and deep learning classification, *IEEE ISCAS*, (2017),1-4, <http://dx.doi.org/10.1109/ISCAS.2017.8050762>.
- [10] Şimşek, E., Özyer, B., Özyer, G. T., Foto-kapan görüntülerinde derin öğrenme tabanlı insan tespiti, *BÜ Fen Bil. Derg.*, 3 (1) (2020), 1-8.
- [11] Şimşek, E., Barış, Ö., Özyer, G. T., Foto-kapan görüntülerinde hareketli nesne tespiti ve konumunun belirlenmesi, *Erzincan University J. Sci. and Tech.*, 12 (2) (2019), 902-919, <https://doi.org/10.18185/erzifbed.509571>.
- [12] ETSI, (2009). Additional spectrum requirements for future public safety and security wireless communication systems in the UHF range. *System Reference Document; Land Mobile Service 1*. Available at: [https://cept.org/files/9421/tr\\_102628v010101p.doc](https://cept.org/files/9421/tr_102628v010101p.doc). [Accessed April 2023].
- [13] Geylani, M., Çibuk, M., Çinar, H., ve Ağgün, F., Geçmişten günümüze hücre sel haberleşme teknolojilerinin gelişimi, *DEÜ Müh. Fak. Fen ve Müh. Derg.*, 18 (54) (2016), 606-623, <http://dx.doi.org/10.21205/deufmd.2016185425>.
- [14] Qaddus, A., Real time performance analysis of Digital Mobile Radio (DMR) and APCO project 25 (P-25) radio systems in Land Mobile Radio (LMR) systems, *Int. J. Comput. Eng. Inf. Tech.*, 8 (3) (2016), 49.
- [15] Savunma Sanayi Başkanlığı (SSB), (2016). Sektör raporu. Available at: [https://thinktech.stm.com.tr/uploads/docs/1608890536\\_stm-sektor-raporu-kamu-guvenligi-ve-acil-yardim.pdf](https://thinktech.stm.com.tr/uploads/docs/1608890536_stm-sektor-raporu-kamu-guvenligi-ve-acil-yardim.pdf). [Accessed March 2023].
- [16] Şahin, A. (2023). *Depremde herşey sustu Aselsan JEMUS konuştu*. Available at: <https://www.savunmasanayist.com/depremde-her-sey-sustu-aselsan-jemus-konustu/>. [Accessed September 2023].
- [17] Şahin, F., Telsiz haberleşme standartları, *İstanbul Aydın Üniversitesi Dergisi*, 27 (2015), (15-30), <http://dx.doi.org/10.17932/IAU.IAUD.m.13091352.2015.7/27.15-30>.
- [18] Babel, L., Coordinated target assignment and UAV path planning with timing constraints, *J. Intell. Robot. Syst.*, 94 (3-4) (2019), 857-869, <http://dx.doi.org/10.1007/s10846-018-0910-9>.
- [19] Koslowski, R., Schulzke, M., Drones along borders: Border security UAVs in the United States and the European Union, *Int. Stud. Perspect.*, 19 (4) (2018), 305-324, <http://dx.doi.org/10.1093/isp/eky002>.
- [20] Haddal, C. C., Gertler, J., Homeland security: Unmanned aerial vehicles and border surveillance, (2010).
- [21] Yildiz, B., Exploration of the use of unmanned aerial vehicles along with other assets to enhance border protection (Doctoral dissertation, Monterey, California. Naval Postgraduate School) (2009).

- [22] Csernatoni, R., Constructing the EU's high-tech borders: FRONTEX and dual-use drones for border management, *European Security*, 27 (2) (2018), 175-200, <http://dx.doi.org/10.1080/09662839.2018.1481396>.
- [23] Villi, O., Yakar, M., İnsansız hava araçlarının kullanım alanları ve sensör tipleri, *TİHA Dergisi*, 4 (2) (2022), 73-100, <http://dx.doi.org/10.51534/tiha.1189263>.
- [24] Arya, L., Rastogi, R., Study on aerial monitoring system in agriculture, forestry, defense, and border protection using artificial intelligence (AI), *Agric. Aquacult. Appl. Biosens. Bioelectron.*, (2024), 389-404, <http://dx.doi.org/10.4018/979-8-3693-2069-3.ch021>.
- [25] Şahiner, M. K., Ayhan, E. and Önder, M., Yeni sınır güvenliği anlayışında yapay zekâ yönetişi: Fırsatlar ve tehditler, *Ulusa*, 5 (2) (2021), 83-95.
- [26] Bakır, G., İnsansız hava araçlarının savunma sanayi harcamasında yeri ve önemi, *ASEAD*, 6 (2) (2019), 127-134, <http://dx.doi.org/10.51534/tiha.884468>.
- [27] Newell, B. C., Gomez, R., Guajardo, V., Sensors, cameras, and the new 'normal' in clandestine migration: How undocumented migrants experience surveillance at the US-Mexico border, *Surveillance and Society*, 15 (1) (2017), 21-41, <http://dx.doi.org/10.24908/ss.v15i1.5604>.
- [28] Berrahal, S., Kim, J. H., Rekhis, S., Boudriga, N., Wilkins, D., Acevedo, J., Border surveillance monitoring using quadcopter UAV-aided wireless sensor networks, *J. Commun. Softw. Syst.*, 12 (1) (2016), 67-82, <http://dx.doi.org/10.24138/jcomss.v12i1.92>.
- [29] Bahaghighat, M., Motamedi, S. A., Xin, Q., Image transmission over cognitive radio networks for smart grid applications, *Applied Sciences*, 9 (24) (2019), 5498, <http://dx.doi.org/10.1109/MWC.2013.6590059>.
- [30] Grois, D., Marpe, D., Mulayoff, A., Itzhaky, B., Hadar, O., Performance comparison of h. 265/mpeg-hevc, vp9, and h. 264/mpeg-avc encoders, *Picture Coding Symposium (PCS) IEEE*, (2013), 394-397.
- [31] Xiong, W., Lv, Y., Zhang, X., Cui, Y., Learning to translate for cross-source remote sensing image retrieval, *IEEE Trans. Geosci. Remote Sens.*, 58 (7) (2020), 4860-4874, <http://dx.doi.org/10.1109/TGRS.2020.2968096>.
- [32] Glass, S., Muthukumarasamy, V., Portmann, M., A software-defined radio receiver for APCO Project 25 signals, *Proceedings of the 2009 ICWCMC*, (2009), 67-72, <http://dx.doi.org/10.1145/1582379.1582395>.
- [33] Ramsey, E. R., A software based APCO Project 25 data transmission base station for local police headquarters, *University of New Hampshire*, (2007), <http://dx.doi.org/10.1109/THS.2008.4534487>.
- [34] Khayam, S. A., The discrete cosine transform (DCT): Theory and application, *Michigan State University*, 114 (1) (2003), 31.
- [35] Ahmed, N., Natarajan, T., Rao, K. R., Discrete cosine transform, *IEEE Trans. on Comp.*, 100 (1) (1974), 90-93.
- [36] Scribano, C., Franchini, G., Prato, M., Bertogna, M., DCT-Former: Efficient self-attention with discrete cosine transform, *J. Sci. Comput.*, 94 (3) (2023), 67, <http://dx.doi.org/10.1007/s10915-023-02125-5>.



## SEARCH POTENTIAL OF THE HIGH ENERGY-LARGE HADRON COLLIDER FOR SPIN-1/2 EXCITED QUARKS IN DI-JET FINAL STATE

Yusuf Oguzhan GÜNAYDIN<sup>1</sup>, Mehmet SAHİN<sup>2</sup> and Leyla AYDIN<sup>1</sup>

<sup>1</sup>Department of Physics, Kahramanmaraş Sütçü İmam University,  
Kahramanmaraş, TÜRKİYE

<sup>2</sup>Department of Computer Engineering, Usak University, Usak, TÜRKİYE

**ABSTRACT.** Composite models, which suggest a possible substructure of fundamental particles, can be directly proven by the discovery of the excited quark. Higher energy and higher-luminosity particle colliders are needed to discover the composite structure predicted in the proposed models. The High Energy Large Hadron Collider (HE-LHC) has the potential to be a possible discovery machine for composite models. In this collider, with a center-of-mass energy of 27 TeV and integrated luminosity between 750 and 15000 fb<sup>-1</sup>, we calculated the exclusion, observation, and discovery limits for the mass of spin-1/2 excited quark in the *di-jet* final state, as well as the attainable compositeness scale values. In addition to these calculations, we scanned free parameters from 0.06 to 1 to determine the HE-LHC potential to reveal spin-1/2 excited quark.

### 1. INTRODUCTION

The Standard Model (SM) is a theory that best describes the interactions between fundamental particles and largely explains the dynamics of these interactions. However, SM cannot provide a sufficient answer to problems such as hierarchy problems, number of families, parameter excess, matter-antimatter asymmetry, quark-lepton symmetry, repetition of fermions, neutrino oscillations, and dark matter. For these problems that particle physicists are trying to solve, new models called Beyond the Standard Model (BSM) theories have emerged. Composite Models are another area of research that predicts the possibility of a substructure of fermions among BSM theories. As research publications that form the basis of Composite Models,

---

*Keywords.* Excited quark, HE-LHC, di-jet, particle phenomenology, composite models, compositeness scale.

✉ yusufgunaydin@gmail.com; 0000-0002-0514-6936

✉ mehmet.sahin@usak.edu.tr-Corresponding author; 0000-0001-6777-3938

✉ leylaydiin@gmail.com; 0000-0001-5562-4265.

we can firstly list an article by Low [1] containing the prediction of heavy electrons and muons and the two papers published later by Jogesh C. Pati and Abdus Salam [2,3]. In the last two publications mentioned, they predicted the composite substructure of fermions, which they called "preons." After these groundbreaking studies, extensive research has been conducted in the literature on the compositeness of fermions and bosons [4–15]. Based on these predictions, many researchers have conducted experimental [16–27] and phenomenological [28–53] studies to discover excited fermions, which will directly prove compositeness.

Researchers hypothesize that excited fermions comprise two distinct elementary particle systems, excited quarks ( $q^*$ ) and excited leptons ( $l^*$ ), similar to SM fermions. The excited quark can exist in four distinct final states: such as *di-jet* ( $q^* \rightarrow jj$ ), *photon-jet* ( $q^* \rightarrow \gamma j$ ), *W-jet* ( $q^* \rightarrow Wj$ ), and *Z-jet* ( $q^* \rightarrow Zj$ ). Experimentally, some exclusion limits have been imposed on the mass of the excited quark for each final state in the CMS and ATLAS experiments at the European Council for Nuclear Research (CERN) with a center-of-mass energy of 13 TeV and total luminosity values of  $35.9 \text{ fb}^{-1}$  and  $139 \text{ fb}^{-1}$ , respectively. These experiments' research established mass limits of 6.7 TeV in the *di-jet* final state, 5.5 TeV in the *photon-jet* final state, 5.0 TeV in the *W-jet* final state, and 4.7 TeV in the *Z-jet* final state [27,54,55].

In-depth exploration of BSM theories necessitates particle colliders characterized by elevated center-of-mass energy and exceptionally high integrated luminosity values. CERN plans to establish the HE-LHC in the 2030s with 27 TeV center-of-mass-energy. This new-generation particle collider can provide a comprehensive spectrum for researching excited quarks, boosting a maximum integrated luminosity of  $15000 \text{ fb}^{-1}$  [56].

The research subject of this study is the discovery ( $5\sigma$ ), observation ( $3\sigma$ ), and exclusion ( $2\sigma$ ) potential of the spin-1/2 excited quark, which transitions to the *di-jet* final state in the HE-LHC using the effective Lagrangian method. In the subsequent sections, we present the interaction Lagrangian, decay widths, cross-section plots, and signal-background analyses of the spin-1/2 excited quark. In the following section, we describe calculations on the discovery, observation, and exclusion mass limits of the spin-1/2 excited quark in the *di-jet* final state at the HE-LHC. Additionally, we discuss the attainable compositeness scale, a crucial parameter in compositeness studies. Furthermore, we analyzed the impact of the spin-1/2 excited quark on the discovery, observation, and exclusion limits by systematically scanning the free parameters that the precise numerical values are unknown. In the last part, the findings are interpreted and discussed.

## 2. MATERIALS AND METHODS

**2.1. Interaction Lagrangian.** In the numerical calculations, we utilized the LanHEP [57] software to incorporate the effective Lagrangian of the spin-1/2 excited quark [47,58] (Equation 1) into the CalcHEP [59] simulation software. Using the

simulation software, we calculated the decay width and cross-section values for the spin-1/2 excited quark in the *di-jet* final state.

$$L_{eff} = \frac{1}{2\Lambda} \overline{q_R^*} \sigma^{\mu\nu} [g_s f_s \frac{\lambda_a}{2} G_{\mu\nu}^a + g f \frac{\vec{\tau}}{2} \vec{W}_{\mu\nu} + g' f' \frac{\Upsilon}{2} B_{\mu\nu}] q_L + h.c. \quad (1)$$

In Equation 1,  $\Lambda$  represents the compositeness scale,  $q_R^*$  represents the right-handed excited quark, and  $q_L$  represents the left-handed SM quark. In addition, the symbols  $g$ ,  $g_s$ , and  $g'$  represent the gauge coupling constants, and the field strength tensors SU(3), SU(2), and U(1) are represented by the symbols  $G_{\mu\nu}^a$ ,  $\vec{W}_{\mu\nu}$ , and  $B_{\mu\nu}$ , respectively. The remaining parameters are expressed as Gell-Mann matrices  $\lambda_a$ , Pauli spin matrices  $\vec{\tau}$ , weak hyper-charge  $\Upsilon$ , and dimensionless free parameters  $f_s$ ,  $f$  and  $f'$ .

**2.2. Decay Widths and Cross Sections.** Excited quarks may consist of three families, like the SM quarks. The parton distribution functions of the  $u$  quark and gluon inside the proton are more dominant than those of other quarks, so in proton-proton colliders, the most dominant production process of excited quarks occurs as  $gu \rightarrow u^* \rightarrow gu$ . As a result, first-family excited quarks will have a higher production cross-section. In contrast, the production cross sections of the second and third family excited quarks, which can be produced in the proton-proton collider through the SM quark-gluon-excited quark vertices corresponding to their respective families, will be much smaller. As an exception to these statements, it can be shown that excited quarks can make transitions between families through the Flavor Changing Neutral Current (FCNC) interactions. Experimentally, it is evident that an FCNC interaction between SM quarks has not been observed at the tree level. Excited quarks entering the FCNC interactions with the SM quarks can cause more complex phenomenological final states. For these reasons, second and third-family excited quark productions and the FCNC interactions of excited quarks with the SM quarks are excluded from the scope of this study. So, this study focused exclusively on examining excited quarks that interact with the first-family quarks of the SM. For excited quarks, (1) excited quark  $u^*$  if  $m_{u^*} > m_{d^*}$ , (2) excited quark  $d^*$  if  $m_{d^*} > m_{u^*}$ , and (3) excited quark  $q^*$  states if  $m_{u^*} = m_{d^*}$  (degenerate) were investigated. In addition, (a)  $\Lambda = 27$  TeV and (b)  $\Lambda = m_{Q^*}$  values were used in the compositeness scale we used in our calculations ( $Q^* : u^*, d^*, \text{ and } q^*$ ). We performed decay width and cross-section calculations, considering the experimentally determined exclusion mass limit of 6.7 TeV for the excited quark in the *di-jet* final state by selecting free parameters as  $f = f_s = f' = 1$ . For the (a) and (b) preferences of the compositeness scale, the decay width graphs of the excited  $d$  quark ( $d^*$ ) and the excited  $u$  quark ( $u^*$ ) contributed by the total and four different channels separately are given in Figures 1 and 2, respectively. In these plots, we observe that the *di-jet* final state contributes the most to the decay widths of the excited  $u$  and  $d$  quarks.

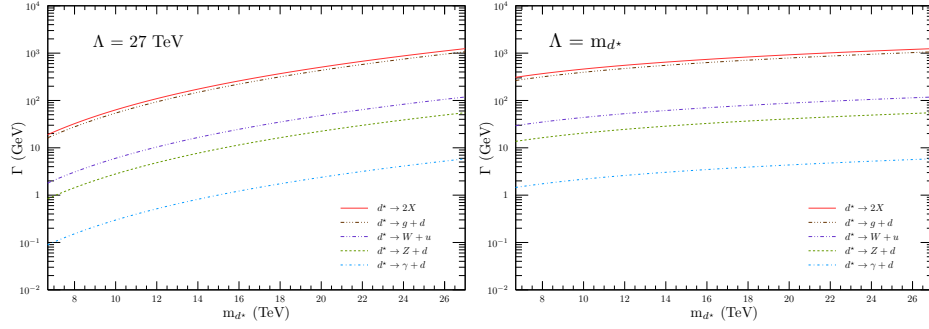


FIGURE 1. Total decay width as a function of spin-1/2 excited  $d$  quark for  $\Lambda = 27$  TeV (left panel) and  $\Lambda = m_{d^*}$  (right panel).

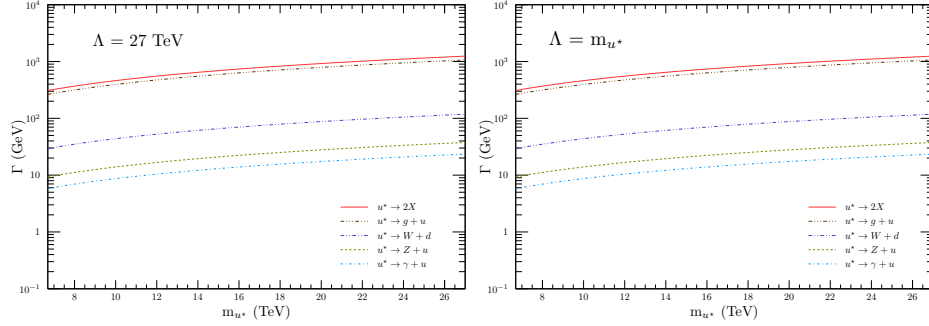


FIGURE 2. Total decay width as a function of spin-1/2 excited  $u$  quark for  $\Lambda = 27$  TeV (left panel) and  $\Lambda = m_{u^*}$  (right panel).

To mitigate potential divergences in the cross-section values, we present the cross-section graphs for  $d^*$ ,  $u^*$ , and  $q^*$  by imposing the constraint of  $P_{T_j} > 25$  GeV, as illustrated in Figure 3. We selected CT10 for the quark distribution function [60], and in these cross-section calculations, we set the renormalization and factorization scales equal to the mass of the excited quarks. Both plots in Figure 3 depict the cross-section values capable of yielding at least one event, starting from the experimental exclusion value of 6.7 TeV aligned with the mass of excited quarks. This consideration incorporates the maximum anticipated integrated luminosity value for the HE-LHC. The disparity between the two plots arises from employing two distinct values of the compositeness scale. As anticipated, setting the compositeness scale equal to the mass of the excited quark results in a higher cross-section.

**2.3. Signal and Background Analysis.** For the signal,  $pp \rightarrow u^* + X \rightarrow ug + X$ ,  $pp \rightarrow d^* + X \rightarrow dg + X$ , and  $pp \rightarrow q^* + X \rightarrow qq + X$  processes were examined

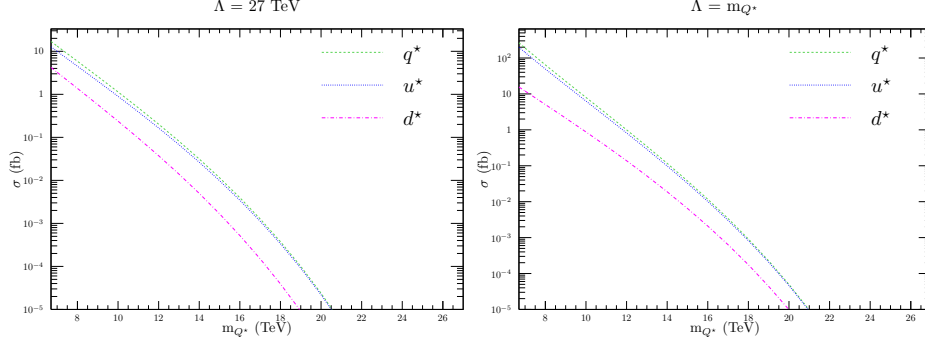


FIGURE 3. Distribution plots in cross-section versus mass for the *di-jet* final state of spin-1/2 excited quarks at  $\Lambda = 27$  TeV (left panel) and  $\Lambda = m_{Q^*}$  (right panel).

separately. As background, the  $pp \rightarrow jj+X$  process meets the mentioned signal processes. We defined the  $j$  symbol in the background process as  $u, \bar{u}, d, \bar{d}, c, \bar{c}, s, \bar{s}, b, \bar{b}$ , and  $g$ , and performed calculations accordingly in the simulation software. Initially, we imposed a transverse momentum limitation of 25 GeV for the jets in the signals and background calculations. However, under this constraint, it became challenging to distinguish between the signal and background. By obtaining and examining the transverse momentum ( $P_{T_j}$ ), pseudo-rapidity ( $\eta_j$ ), and invariant mass ( $m_{jj}$ ) distributions, we determined the necessary limitations for these three important parameters that we will use in our later calculations. In Figure 2.3, we present distribution plots for  $P_{T_j}$ ,  $\eta_j$ , and  $m_{jj}$ , featuring only the case where the compositeness scale is equal to the mass of the excited quark ( $\Lambda = m_{Q^*}$ ), as the  $\Lambda = 27$  TeV scenario exhibits a comparable distribution.

Examining the distributions in Figure 2.3, we define the constraints applied in our calculations for  $P_{T_j}$ ,  $\eta_j$ , and  $m_{jj}$ . While determining the  $P_{T_j}$  cut from the plot, the value at which the background is suppressed, and the signal unaffected was selected as 2 TeV. It can be seen in the  $\eta_j$  plot that in the cut applied by choosing between -2.5 and 2.5, a large part of the background will be suppressed. When looking at the  $m_{jj}$  plot, the region within  $m_{Q^*} - 2\Gamma^* < m_{jj} < m_{Q^*} + 2\Gamma^*$  of the peaks where the signal is higher than the background was selected as an invariant mass cut. Here,  $m_{Q^*}$  represents the individual masses of all degenerate and non-degenerate excited quarks, and  $\Gamma^*$  represents the decay width of excited quarks.

In addition to these three critical constraints, we selected a cone angle radius of  $\Delta R > 0.5$  to enhance the distinction of jets in the *di-jet* final state. Utilizing the specified constraints, we computed the discovery, observation, and exclusion limits on the excited quark mass by the Statistical Significance (SS) relation outlined in



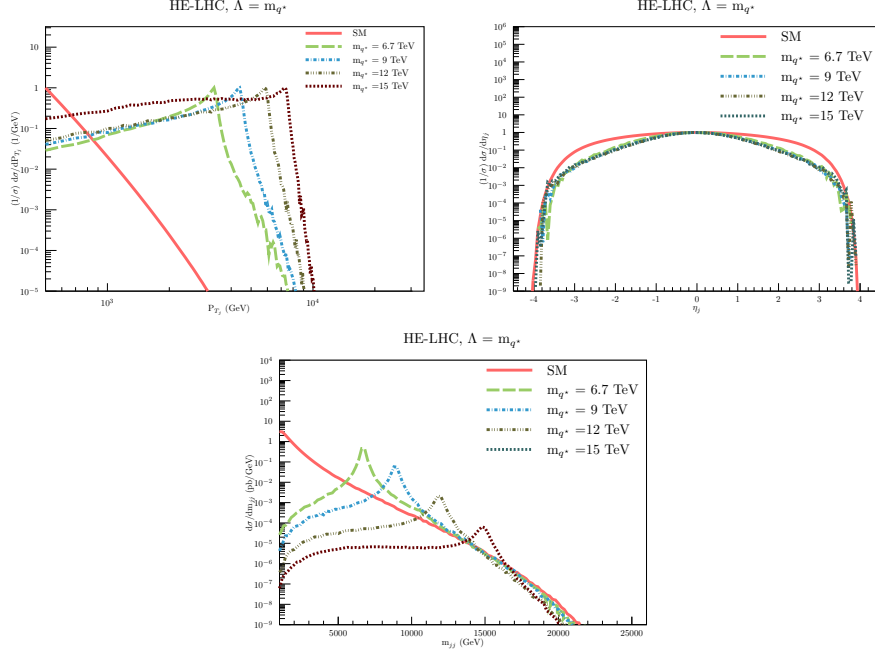


FIGURE 4. Normalized transverse momentum, normalized pseudo-rapidity, and invariant mass distributions of the *di-jet* final-state excited quark for some mass values at HE-LHC.

Equation 2. In this relation,  $\sigma_S$  symbolizes the signal cross section,  $\sigma_B$  symbolizes the background cross section, and  $\mathcal{L}_{int}$  represents the integrated luminosity value.

$$SS = \frac{\sigma_S}{\sqrt{\sigma_S + \sigma_B}} \sqrt{\mathcal{L}_{int}} \quad (2)$$

### 3. FINDINGS AND CONCLUSIONS

To reveal the ability of the HE-LHC to investigate the spin-1/2 excited quark, firstly, using the cross-section results obtained with the help of CalcHEP simulation software and the statistical significance relation in Equation 2, the spin-1/2 excited quark mass limits were calculated considering the all confidence level  $2\sigma$  (exclusion),  $3\sigma$  (observation), and  $5\sigma$  (discovery). In Table 1, we consider the compositeness scale as 27 TeV. We set the integrated luminosity value for the first year of HE-LHC at  $750 \text{ fb}^{-1}$ , and we utilize the projected integrated luminosity value of 15000  $\text{fb}^{-1}$  for the end of 20 years. The table displays the discovery, observation, and exclusion mass limits for the *di-jet* final state of  $d^*$ ,  $u^*$  and  $q^*$ . As expected, the mass limits of the spin-1/2 excited quark in the degenerate state were higher. In

addition, we observe that the potential exclusion limit for the spin-1/2  $d^*$  *di-jet* final state, which possesses the lowest mass limit in our calculations, is anticipated to significantly surpass the exclusion limit of 6.7 TeV set by the LHC in the first year. As the integrated luminosity value increases, one anticipates that these mass limits will achieve higher values. Although Table 1 numerically presents the statistical significance values when the integrated luminosity reaches 15000 fb<sup>-1</sup>, Figure 5 illustrates the 20-year developmental trajectory of these mass limits.

TABLE 1. Exclusion, observation, and discovery mass limits obtained with the lowest and highest integrated luminosity values of the HE-LHC for the case where the compositeness scale  $\Lambda$  is taken equal to 27 TeV.

$\mathcal{L}_{int}$ (fb <sup>-1</sup> ) :	750			15000		
SS* :	2 $\sigma$	3 $\sigma$	5 $\sigma$	2 $\sigma$	3 $\sigma$	5 $\sigma$
$m_{d^*}$ (TeV) :	10.3	9.7	8.8	12.6	12.0	11.2
$m_{u^*}$ (TeV) :	12.6	11.9	11.1	15.1	14.5	13.6
$m_{q^*}$ (TeV) :	12.9	12.1	11.5	15.4	14.7	13.9

\*SS = Statistical Significance

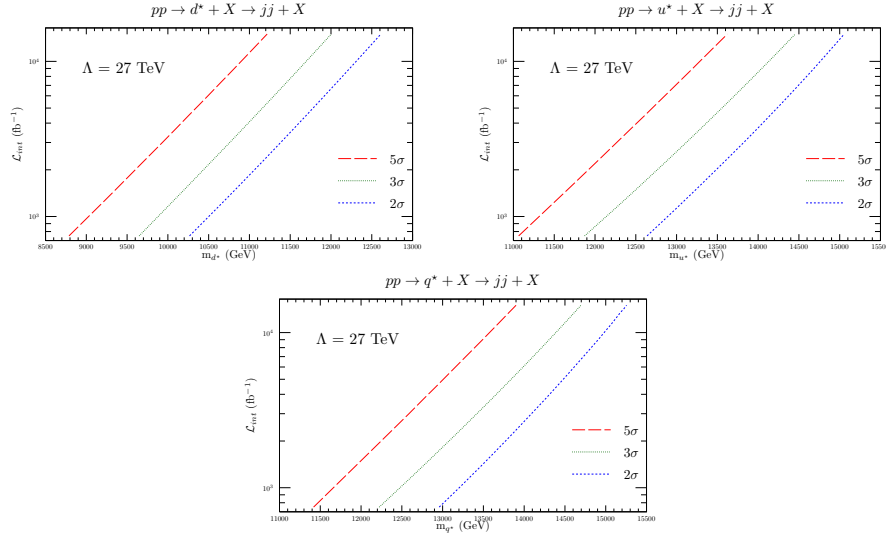


FIGURE 5. Discovery, observation, and exclusion mass limits of the spin-1/2 excited  $d$ ,  $u$  and  $q$  quarks in the *di-jet* final state according to the integrated luminosity values of the HE-LHC for  $\Lambda = 27$  TeV.

We conducted an additional calculation to determine the mass limit of spin-1/2 excited  $d$ ,  $u$ , and  $q$  quarks in the *di-jet* final state at the HE-LHC while considering

the compositeness scale taken equal to the mass of the excited quark. The result of our calculations here is higher than the discovery, observation, and exclusion values in the case of  $\Lambda = 27$  TeV, as seen in Table 2. Based on the results obtained at this juncture, it is evident that the HE-LHC could achieve an exclusion value significantly surpassing the experimentally imposed exclusion limit on the mass of the excited quark, even within its inaugural year of operation, as seen in Figure 6. These calculations underscore the high potential of the HE-LHC for the discovery of excited quarks.

TABLE 2. Exclusion, observation, and discovery mass limits obtained with the lowest and highest integrated luminosity values of the HE-LHC for the case where the compositeness scale  $\Lambda$  is taken equal to the mass of the excited quark.

$\mathcal{L}_{int}$ (fb $^{-1}$ ) :	750			15000		
SS*	2 $\sigma$	3 $\sigma$	5 $\sigma$	2 $\sigma$	3 $\sigma$	5 $\sigma$
$m_{d^*}$ (TeV) :	11.7	11.1	10.4	13.6	13.1	12.4
$m_{u^*}$ (TeV) :	13.8	13.2	12.4	15.9	15.3	14.6
$m_{q^*}$ (TeV) :	14.0	13.5	12.7	16.0	15.5	14.8

\*SS = Statistical Significance

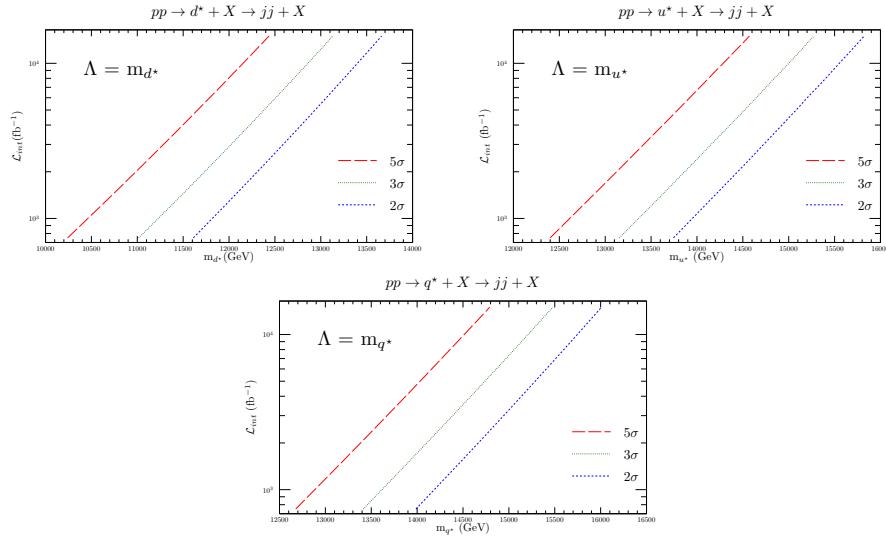


FIGURE 6. Discovery, observation, and exclusion mass limits of the spin-1/2 excited d, u and q quarks in the *di-jet* final state according to the integrated luminosity values of the HE-LHC for  $\Lambda = m_{Q^*}$ .

In our analysis thus far, we have equated the compositeness scale, a pivotal parameter in compositeness studies, either to the center-of-mass energy of the particle collider or the mass of the spin-1/2 excited quark. Nevertheless, the compositeness scale remains an indeterminate parameter. To address this issue, we conducted independent calculations to determine the potential compositeness scale limits achievable in the HE-LHC regarding the masses of the spin-1/2 excited quarks in the *di-jet* final state, specifically  $d^*$ ,  $u^*$ , and  $q^*$ . Table 3 details our compositeness scale calculations, considering the highest value of the integrated luminosity at HE-LHC, which is  $15000 \text{ fb}^{-1}$ . We present achievable compositeness scale values corresponding to exclusion, observation, and discovery limits at selected mass values, such as 6.7, 8.7, 10.7, and 12.7 TeV for spin-1/2  $d^*$ ,  $u^*$ , and  $q^*$  in the *di-jet* final state. At the end of 20 years, when the integrated luminosity value that the HE-LHC will achieve is  $15000 \text{ fb}^{-1}$ , the exclusion values for the compositeness scale corresponding to masses of 6.7 TeV for  $d^*$ ,  $u^*$ , and  $q^*$  will be 920, 2500, and 3200 TeV, respectively. Figure 7 depicts the compositeness scale scan corresponding to the mass of the *di-jet* final state spin-1/2 excited d, u, and q quarks. Our analysis, initiated from the 6.7 TeV exclusion limit imposed by the LHC on the mass of the excited quark, reveals the potential of the HE-LHC to explore the *di-jet* final state excited quark. Considering our findings, it is conceivable that the HE-LHC might function as a potential discovery machine for excited quarks.

TABLE 3. Achievable compositeness scale values corresponding to some mass values of the excited quark at  $15000 \text{ fb}^{-1}$  luminosity of the HE-LHC with center-of-mass energy of 27 TeV.

HE-LHC ( $\mathcal{L}_{int}=15000 \text{ fb}^{-1}$ )									
$m_{Q^*}$ (TeV)	$\Lambda$ (TeV)								
	$m_{d^*}$			$m_{u^*}$			$m_{q^*}$		
	$5\sigma$	$3\sigma$	$2\sigma$	$5\sigma$	$3\sigma$	$2\sigma$	$5\sigma$	$3\sigma$	$2\sigma$
6.7	368	614	920	1000	1668	2500	1280	2133	3200
8.7	126	210	314	400	666	999	503	839	1258
10.7	38.4	64	96	145	242	362	179	298	447
12.7	11	18	27	48	80	120	58	96	144

In the calculations thus far, we have set the free parameters to equal one. The value  $f = f_s = f' = 1$  is the highest value that free parameters can take, but these parameters can also have values between zero and one. Thus, scanning the numeric value of the free parameters becomes beneficial through the mass of the excited quark. Given that the most extreme value attained in the free parameter scan pertains to the scan for excited q quarks, the results for  $d^*$  and  $u^*$  scans are not incorporated into this study. Figure 8 shows the scans we made considering  $\Lambda = 27 \text{ TeV}$  and  $\Lambda = m_{q^*}$  according to the mass of the excited q quark. In the calculations involving the scanning of free parameters, we employed an integrated

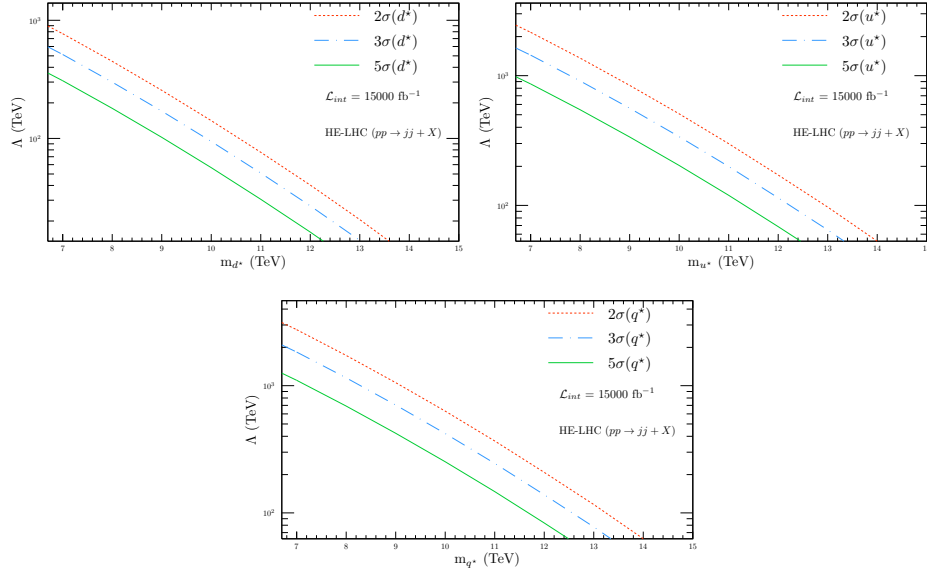


FIGURE 7. Achievable compositeness scale values corresponding to the discovery, observation, and exclusion masses of  $d^*$ ,  $u^*$  and  $q^*$  when the integrated luminosity value of HE-LHC is  $15000 \text{ fb}^{-1}$ .

luminosity value of  $15000 \text{ fb}^{-1}$ , representing the HE-LHC can attain over 20 years. Upon examination of the scan plot for  $\Lambda = 27 \text{ TeV}$ , the discovery, observation, and exclusion limits of the excited  $q$  quark are apparent, being  $6.7 \text{ TeV}$ ,  $8.0 \text{ TeV}$ , and  $8.8 \text{ TeV}$ , respectively, when the free parameters are chosen to  $0.13$ . For the scenario where  $\Lambda$  equals the mass of the excited quark ( $\Lambda = m_{q^*}$ ) with the free parameters set to  $0.06$ , the discovery, observation, and exclusion limits of the excited  $q$  quark are identified to be  $6.7 \text{ TeV}$ ,  $7.5 \text{ TeV}$ , and  $8.2 \text{ TeV}$ , respectively. As depicted in Figure 8, the numerical increase in the value of the free parameters results in an increase in the discovery, observation, and exclusion limits placed on the mass of the excited  $q$  quark. The calculations, conducted through a scan of the free parameters and unveiling the potential for values as small as  $0.06$ , indicate the considerable research potential of the HE-LHC for the spin-1/2 excited quark in the *di-jet* final state.

This study encompasses three distinct analyses: determination of discovery, observation, and exclusion limits on the mass of spin-1/2 excited quarks in the *di-jet* final state; assessment of achievable compositeness scale; and screening of free parameters in the context of the High Energy-Large Hadron Collider, taking into account two separate integrated luminosity values. The outcomes of these analyses reveal that the HE-LHC exhibits a significantly greater potential than the LHC in

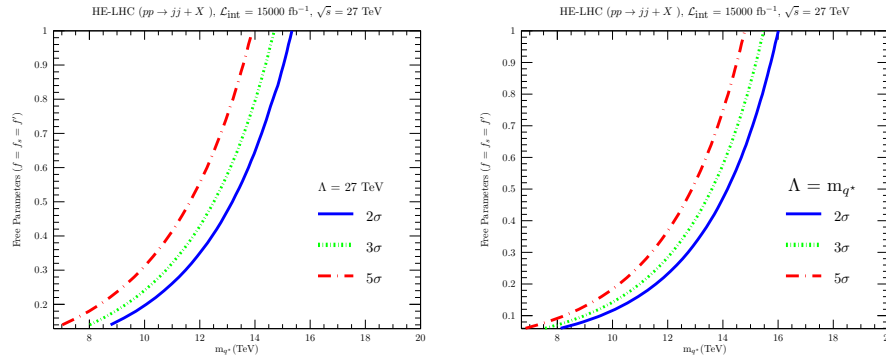


FIGURE 8. In the cases  $\Lambda = 27$  TeV (left panel) and  $\Lambda = m_{q^*}$  (right panel), the attainable mass of  $q^*$  according to the free parameter values limits.

achieving higher limits for the mass of excited quarks, larger values for achievable compositeness scale, and distinctly small values for free parameters.

**Author Contribution Statements** Authors are equally contributed to the paper. All authors read and approved the final copy of the manuscript.

**Declaration of Competing Interests** The authors did not receive support from any organization for the submitted work. The authors have no relevant financial or non-financial interests to disclose.

**Acknowledgement** We thank Usak University, Energy, Environment and Sustainability Application and Research Center for supporting this study.

## REFERENCES

- [1] Low, F. E., Heavy electrons and muons, *Phys. Rev. Lett.*, 14 (7) (1965), 238–239, <https://doi.org/10.1103/PhysRevLett.14.238>.
- [2] Pati, Jogesh C. and Salam, A., Lepton number as the fourth “color”, *Phys. Rev. D*, 10 (1) (1974), 275–289, <https://doi.org/10.1103/PhysRevD.10.275>.
- [3] Pati, J. C., Salam A., and Strathdee, J., Are quarks composite?, *Phys. Lett. B*, 59 (3) (1975), 265–268, [https://doi.org/10.1016/0370-2693\(75\)90042-8](https://doi.org/10.1016/0370-2693(75)90042-8).
- [4] Terazawa, H., Chikashige, Y., and Akama K., Unified model of the nambu-jona-lasinio type for all elementary-particle forces., *Phys. Rev. D*, 15 (2) (1977), 480–487, <https://doi.org/10.1103/PhysRevD.15.480>.
- [5] Shupe, M. A., A composite model of leptons and quarks., *Phys. Lett. B*, 86 (1979), 87–92, [https://doi.org/10.1016/0370-2693\(79\)90627-0](https://doi.org/10.1016/0370-2693(79)90627-0).
- [6] Harari, H., A schematic model of quarks and leptons., *Phys. Lett. B*, 86 (1) (1979), 83–86, [https://doi.org/10.1016/0370-2693\(79\)90626-9](https://doi.org/10.1016/0370-2693(79)90626-9).
- [7] Fritzsch, H. and Mandelbaum G., Weak-interactions as manifestations of the substructure of leptons and quarks., *Phys. Lett. B*, 102 (5), (1981), 319–322, [https://doi.org/10.1016/0370-2693\(81\)90626-2](https://doi.org/10.1016/0370-2693(81)90626-2).

- [8] Terazawa, H., A fundamental theory of composite-particles and fields., *Phys. Lett. B*, 133 (1-2) (1983), 57–60, [https://doi.org/10.1016/0370-2693\(83\)90105-3](https://doi.org/10.1016/0370-2693(83)90105-3).
- [9] Eichten, E. J., Lane, K. D., and Peskin, M. E., New tests for quark and lepton substructure., *Phys. Rev. Lett.*, 50 (11) (1983), 811–814, <https://doi.org/10.1103/PhysRevLett.50.811>.
- [10] D'Souza, I. A. and Kalman, C. S., *Preons: Models of leptons, quarks and gauge bosons as composite objects*, (1992).
- [11] Çelikel, A., Kantar, M., and Sultansoy, S., A search for sextet quarks and leptogluons at the LHC, *Phys. Lett. B*, 443 (1-4) (1998), 359–364, [https://doi.org/10.1016/S0370-2693\(98\)01299-4](https://doi.org/10.1016/S0370-2693(98)01299-4).
- [12] De Souza, M. E., Weak decays of hadrons reveal compositeness of quarks., *Sci. Ple.*, (6), (2008), <https://www.scienciaplenua.org.br/sp/article/view/612>.
- [13] Terazawa, H. and Yasue, M., Excited gauge and higgs bosons in the unified composite model., *Nonlin. Phenom. Complex Syst.*, 19 (1) (2016), 1–6.
- [14] Fritzsche, H., Composite weak bosons at the large hadronic collider., *Mod. Phys. Lett. A*, 31 (20) (2016), 1630019, <https://doi.org/10.1142/S0217732316300196>.
- [15] Kaya, U., Oner, B. B., and Sultansoy, S., A minimal fermion-scalar preonic model., *Turkish J. of Phys.*, 42 (3) (2018), 235–241, <https://doi.org/10.3906/fiz-1710-28>.
- [16] Adloff, C. et al., A search for excited fermions at HERA, *Eur. Phys. J. C*, 17 (4) (2000), 567–581, <https://doi.org/10.1007/s100520000503>.
- [17] Acciarri, M. et al., Search for excited leptons in  $e^+e^-$  interactions at  $s=192-202$  GeV, *Phys. Lett. B*, 502 (1-4) (2001), 37–50, [https://doi.org/10.1016/S0370-2693\(01\)00133-2](https://doi.org/10.1016/S0370-2693(01)00133-2).
- [18] Chekanov, S. et al., Searches for excited fermions in ep collisions at HERA, *Phys. Lett. B*, 549 (1-2) (2002), 32–47, [https://doi.org/10.1016/S0370-2693\(02\)02863-0](https://doi.org/10.1016/S0370-2693(02)02863-0).
- [19] Chatrchyan, S. et al., Search for resonances in the dijet mass spectrum from 7 TeV  $pp$  collisions at CMS, *Phys. Lett. B*, 704 (3) (2011), 123–142, <https://doi.org/10.1016/j.physletb.2011.09.015>.
- [20] Khachatryan, V. et al., Search for excited quarks in the  $\gamma + \text{jet}$  final state in proton-proton collisions at  $\sqrt{s} = 8$  TeV, *Phys. Lett. B*, 738 (2014), 274–293, <https://doi.org/10.1016/j.physletb.2014.09.048>.
- [21] Khachatryan, V. et al., Search for Narrow Resonances Decaying to Dijets in Proton-Proton Collisions at  $\sqrt{s} = 13$  TeV, *Phys. Rev. Lett.*, 116 (7) (2016), 071801, <https://doi.org/10.1103/PhysRevLett.116.071801>.
- [22] Aad, G. et al., Search for new phenomena in dijet mass and angular distributions from  $pp$  collisions at  $\sqrt{s}=13$  TeV with the ATLAS detector, *Phys. Lett. B*, 754 (2016), 302–322, <https://doi.org/10.1016/j.physletb.2016.01.032>.
- [23] Aad, G. et al., Search for new phenomena with photon+jet events in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector, *JHEP*, 2016 (3) (2016), 41, [https://doi.org/10.1007/jhep03\(2016\)041](https://doi.org/10.1007/jhep03(2016)041).
- [24] Aaboud, M. et al., Search for new phenomena in dijet events using 37  $\text{fb}^{-1}$  of  $pp$  collision data collected at  $\sqrt{s} = 13$  TeV with the ATLAS detector, *Phys. Rev. D*, 96 (5) (2017), 052004, <https://doi.org/10.1103/PhysRevD.96.052004>.
- [25] Sirunyan, A. M. et al., Search for dijet resonances in proton–proton collisions at  $\sqrt{s} = 13$  TeV and constraints on dark matter and other models, *Phys. Lett.*, B769 (2017), 520–542, <https://doi.org/10.1016/j.physletb.2017.02.012>.
- [26] Sirunyan, A. M. et al., Search for narrow and broad dijet resonances in proton-proton collisions at  $\sqrt{s} = 13$  TeV and constraints on dark matter mediators and other new particles, *JHEP*, 2018 (8) (2018), 130, [https://doi.org/10.1007/jhep08\(2018\)130](https://doi.org/10.1007/jhep08(2018)130).
- [27] Sirunyan, A. M. et al., Search for massive resonances decaying into  $WW$ ,  $WZ$ ,  $ZZ$ ,  $qZ$ , and  $qZ$  with dijet final states at  $\sqrt{s} = 13$  TeV, *Phys. Rev. D*, 97 (7) (2018), 072006, <https://doi.org/10.1103/PhysRevD.97.072006>.

- [28] Renard, F. M., Excited quarks and new hadronic states, *Il Nuovo Cimento*, 77 (1) (1983), 1–20.
- [29] Lyons, L., An introduction to the possible substructure of quarks and leptons. *Prog. Part. Nucl. Phys.*, 10 (1983), 227–304, [https://doi.org/10.1016/0146-6410\(83\)90005-4](https://doi.org/10.1016/0146-6410(83)90005-4).
- [30] Kuhn, J. and Zerwas, P., Excited quarks and leptons. *Phys. Lett. B*, 147 (1-3) (1984), 189–196, [https://doi.org/10.1016/0370-2693\(84\)90618-X](https://doi.org/10.1016/0370-2693(84)90618-X).
- [31] Pancheri, G. and Srivastava, Y. N., Weak isospin spectroscopy of excited quarks and leptons. *Phys. Lett. B*, 146 (1-2) (1984), 87–94, [https://doi.org/10.1016/0370-2693\(84\)90649-X](https://doi.org/10.1016/0370-2693(84)90649-X).
- [32] de Rújula, A., Maiani, L., and Petronzio, R., Search for excited quarks, *Phys. Lett. B*, 140 (3-4) (1984), 253–258, [https://doi.org/10.1016/0370-2693\(84\)90930-4](https://doi.org/10.1016/0370-2693(84)90930-4).
- [33] Kuhn, J. H., Tholl, H. D., and Zerwas, P. M., Signals of excited quarks and leptons, *Phys. Lett. B*, 158 (3) (1985), 270–275, [https://doi.org/10.1016/0370-2693\(85\)90969-4](https://doi.org/10.1016/0370-2693(85)90969-4).
- [34] Hagiwara, K., Komamiya, S., and Zeppenfeld, D., Excited lepton production at lep and hera. *Z. Phys. C*, 29 (1) (1985), 115–122, <https://doi.org/10.1007/Bf01571391>.
- [35] Baur, U., Hinchliffe, I., and Zeppenfeld, D., Excited quark production at hadron colliders. *Int. J. Mod. Phys. A*, 02 (04) (1987), 1285–1297, <https://doi.org/10.1142/s0217751x87000661>.
- [36] Spira, M. and Zerwas, P. M., *Excited Quarks and Leptons*, pages 519–529, Springer, (1989).
- [37] Jikia, G., Excited quark production at  $ep$  and  $\gamma p$  colliders, *Nucl. Phys. B*, 333 (2) (1990), 317–334, [https://doi.org/10.1016/0550-3213\(90\)90040-k](https://doi.org/10.1016/0550-3213(90)90040-k).
- [38] Baur, U., Spira, M. and Zerwas, P. M., Excited-quark and -lepton production at hadron colliders. *Phys. Rev. D*, 42 (3) (1990), 815–824, <https://doi.org/10.1103/physrevd.42.815>.
- [39] Boudjema, F., Djouadi, A., and Kneur, J. L., Excited fermions at  $e^+e^-$  and  $ep$  colliders. *Z. Phys. C*, 57 (3) (1993), 425–449, <https://doi.org/10.1007/bf01474339>.
- [40] Cakir, O. and Mehdiyev, R., Excited quark production at the CERN LHC. *Phys. Rev. D*, 60 (3) (1999), 034004, <https://doi.org/10.1103/PhysRevD.60.034004>.
- [41] Cakir O., Leror C. and Mehdiyev R., Search for excited quarks with the ATLAS experiment at the CERN LHC: Double jets channel. *Phys. Rev. D*, 62 (11) (2000), 114018, <https://doi.org/10.1103/PhysRevD.62.114018>.
- [42] Cakir, O., Leroy, C., and Mehdiyev, R., Search for excited quarks with the ATLAS experiment at the CERN LHC:  $W/Z + jet$  channel, *Phys. Rev. D*, 63 (9) (2001), 094014, <https://doi.org/10.1103/PhysRevD.63.094014>.
- [43] Eboli, O. J. P., Lietti, S. M., and Mathews, P., Excited leptons at the cern large hadron collider. *Phys. Rev. D*, 65 (7) (2002), 075003, <https://doi.org/10.1103/PhysRevD.65.075003>.
- [44] Cakir, O., Yilmaz, A., and Sultansoy, S., Single production of excited electrons at future  $e^+e^-$ ,  $ep$  and  $pp$  colliders, *Phys. Rev. D*, 70 (7) (2004), 075011, <https://doi.org/10.1103/PhysRevD.70.075011>.
- [45] Cakir, O., Leroy, C., Mehdiyev, R., and Belyaev, A., Production and decay of excited electrons at the LHC. *The Eur. Phys. J. C*, 32 (2) (2004), 1–17, <https://doi.org/10.1140/epjcd/s2003-01-005-5>.
- [46] Cakir, O., and Ozansoy, A., Search for excited spin-3/2 and spin-1/2 leptons at linear colliders. *Phys. Rev. D*, 77 (3) (2008), 035002, <https://doi.org/10.1103/PhysRevD.77.035002>.
- [47] Baur, U., Hinchliffe, I., and Zeppenfeld, D., Excited quark production at hadron colliders, *Int. J. Mod. Phys. A*, 02 (04) (2012), 1285–1297, <https://doi.org/10.1142/s0217751x87000661>.
- [48] Caliskan, A., Excited neutrino search potential of the FCC-based electron-hadron colliders, *Adv. High Energy Phys.*, 2017 (2017), 1–9, <https://doi.org/10.1155/2017/4726050>.
- [49] Caliskan, A. and Kara, S. O., Single production of the excited electrons in the future FCC-based lepton-hadron colliders, *Int. J. Mod. Phys. A*, 33 (24) (2018), 1850141, <https://doi.org/10.1142/S0217751x18501415>.
- [50] Caliskan, A. and Kara, S. O. and Ozansoy, A., Excited muon searches at the FCC-based muon-hadron colliders, *Adv. High Energy Phys.*, 2017 (2017), 1–9, <https://doi.org/10.1155/2017/1540243>.



- [51] Günaydin, Y. O., Sahin, M., and Sultansoy, S., Resonance Production of Excited  $u$  Quark at FCC-based  $\gamma p$  Colliders, *Acta Phys. Polon. B*, 49 (10) (2018), 1763, <https://doi.org/10.5506/APhysPolB.49.1763>.
- [52] Akay, A. N., Günaydin, Y. O., Sahin, M., and Sultansoy, S., Search for Excited  $u$  and  $d$  Quarks in Dijet Final States at Future  $pp$  Colliders, *Adv. High Energy Phys.*, 2019 (2019), 1–11, <https://doi.org/10.1155/2019/9090785>.
- [53] Sahin, M., Aydin, G. and Günaydin, Y. O., Excited quarks production at FCC and SPPC  $pp$  colliders, *Intl J Mod. Phys. A*, 34 (29) (2019), 1950169, <https://doi.org/10.1142/S0217751x19501690>.
- [54] Sirunyan, A. M. et al., Search for excited quarks of light and heavy flavor in  $\gamma + \text{jet}$  final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV, *Phys. Lett. B*, 781 (2018), 390–411, <https://doi.org/10.1016/j.physletb.2018.04.007>.
- [55] Aad, G. et al., Search for new resonances in mass distributions of jet pairs using 139 fb<sup>-1</sup> of  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector, *JHEP*, 2020 (3) (2020), 145, [https://doi.org/10.1007/jhep03\(2020\)145](https://doi.org/10.1007/jhep03(2020)145).
- [56] Abada, A. et al., HE-LHC: The High-Energy Large Hadron Collider Future Circular Collider Conceptual Design Report (Volume 4), *Eur. Phys. J. S.T.*, 228 (5) (2019), 1109–1382, <https://doi.org/10.1140/epjst/e2019-900088-6>.
- [57] Semenov, A., LANCHEP - a package for automatic generation of feynman rules from the lagrangian. version 3.2, *Comp. Phys. Comm.*, 201 (2016), 167–170, <https://doi.org/10.1016/j.cpc.2016.01.003>.
- [58] Zyla, P. A. et al., Review of particle physics, *Prog. Theo. Exp. Phys.*, 2020 (8), (2020), <https://doi.org/10.1093/ptep/ptaa104>.
- [59] Belyaev, A., Christensen N. D., and Pukhov A., CalcHEP 3.4 for collider physics within and beyond the standard model, *Comp. Phys. Comm.*, 184 (7) (2013), 1729–1769, <https://doi.org/10.1016/j.cpc.2013.01.014>.
- [60] Lai, H-L. , Guzzi, M., Huston, J., Li Z., Nadolsky, P. M., Pumplin, J., and Yuan, C. P., New parton distributions for collider physics, *Phys. Rev. D*, 82 (7) (2010), 074024, <https://doi.org/10.1103/PhysRevD.82.074024>.

## INSTRUCTIONS TO CONTRIBUTORS

**Communications Faculty of Sciences University of Ankara Series A2-A3: Physical Sciences and Engineering** is a single-blind peer reviewed open access journal which has been published since 1948 by Ankara University, accepts original research articles written in English in the fields of Physics, Engineering Physics, Electronics/Computer Engineering, Astronomy and Geophysics. Review articles written by eminent scientists can also be invited by the Editor.

**Article-processing charges:** The publication costs for Communications Faculty of Sciences University of Ankara Series A2-A3: Physical Sciences and Engineering are covered by the journal, so authors do not need to pay an article-processing and submission charges. The PDF copies of accepted papers are free of charges and can be downloaded from the website. Hard copies of the paper, if required, are due to be charged for the amount of which is determined by the administration each year.

**Submission:** All manuscripts should be submitted via our online submission: <https://dergipark.org.tr/en/journal/2457/submission/step/manuscript/new> Note that only two submissions per author per year will be considered. Once a paper is submitted to our journal, all co-authors need to wait 6 months from the submission date before submitting another paper.

**Cover Letter:** Manuscripts should be submitted in the PDF form used in the peer-review process together with **THE COVER LETTER** and the source file (Supporting File). In the cover letter the authors should suggest the most appropriate Area Editor for the manuscript and potential four reviewers with full names, universities and institutional email addresses. Proposed reviewers must be experienced researchers in your area of research and at least two of them should be from different countries. In addition, proposed reviewers must not be co-authors, advisors, students, etc. of the authors. In the cover letter, the author may enter the name of anyone who he/she would prefer not to review the manuscript, with detailed explanation of the reason. Note that the editorial office may not use these nominations, but this may help to speed up the selection of appropriate reviewers.

**Preparing your manuscript:** Manuscripts should be typeset using as DOC or LaTeX. Authors will submit their manuscript and the cover letter via our submission system. A template of manuscript can be reviewed in <https://dergipark.org.tr/tr/download/journal-file/20554> (or can be reviewed in [pdf form](#)).

**Title Page:** The title page should contain the title of the paper, full names of the authors, affiliations addresses and e-mail addresses of all authors. Authors are also required to submit their Open Researcher and Contributor ID (ORCID)'s which can be obtained from <http://orcid.org> as their URL address in the format <http://orcid.org/xxxx-xxxx-xxxx-xxxx>. Please indicate the corresponding author.

**Abstract and Keywords:** The abstract should state briefly the purpose of the research. The length of the Abstract should be between 50 to 5000 characters. At least 3 keywords are required.

**Math Formulae:** Formulas should be numbered consecutively in the parentheses ( ).

**Tables:** All tables must have numbers (TABLE 1) consecutively in accordance with their appearance in the text and a legend above the table. Please submit tables as editable text not as images.

**Figures:** All figures must have numbers (FIGURE 1) consecutively in accordance with their appearance in the text and a caption (not on the figure itself) below the figure. Please submit figures as EPS, PDF, TIFF or JPEG format.

Authors Contribution Statement, Declaration of Competing Interests and Acknowledgements should be given at the end of the article before the references.

**References:** The following format for the references should be used. Authors are urged to use the Communication.csl style (<https://dergipark.org.tr/en/download/journal-file/18514>) in Mendeley Desktop or Zotero automated bibliography. If manual entry is preferred for bibliography, then all citations must be listed in the references part and vice versa. Below, It has no relationship with the text, but can be used to show sample citations such as; for articles [1, 4], for books/booklets/theses [3], and for proceedings/conferences etc. [2].

[1] Demirci, E., Unal, A., Özalp, N., A fractional order SEIR model with density dependent death rate, Hacettepe J. Math. Stat., 40 (2) (2011), 287–295.

[2] Gairola, A. R., Deepmala, Mishra, L. N., Rate of approximation by finite iterates of q-Durrmeyer operators, Proc. Natl. Acad. Sci. India Sect. A Phys. Sci., 86 (2) (2016), 229–234.

[3] Lehmann, E. L., Casella, G., Theory of Point Estimation, Springer, New York, 2003.

[4] Özalp, N., Demirci, E., A fractional order SEIR model with vertical transmission, Math. Comput. Model., 54 (1-2) (2011), 1–6, <https://dx.doi.org/10.1016/j.mcm.2010.12.051>.

**Peer-review policy:** The Editor may seek the advice of two, or three referees, depending on the response of the referees, chosen in consultation with appropriate members of the Editorial Board, from among experts in the field of specialization of the paper. The reviewing process is conducted in strict confidence and the identity of a referee is not disclosed to the authors at any point since we use a single-blind peer review process.

**Copyright:** Copyright on any open access article in Communications Faculty of Sciences University of Ankara Series A2-A3: Physical Sciences and Engineering is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

**Declarations/Ethics** With the submission of the manuscript authors declare that:

- All authors of the submitted research paper have directly participated in the planning, execution, or analysis of study;
- All authors of the paper have read and approved the final version submitted;
- The contents of the manuscript have not been submitted, copyrighted or published elsewhere and the visual-graphical materials such as photograph, drawing, picture, and document within the article do not have any copyright issue;
- The contents of the manuscript will not be copyrighted, submitted, or published elsewhere, while acceptance by the Journal is under consideration.
- The article is clean in terms of 'plagiarism', and the legal and ethical responsibility of the article belong to the author(s). Author(s) also accept that the manuscript may go through plagiarism check using iThenticate software;
- The objectivity and transparency in research, and the principles of ethical and professional conduct have been followed. Authors have also declared that they have no potential conflict of interest (financial or non-financial), and their research does not involve any human participants and/or animals.

**Archiving:** Research papers published in Communications Faculty of Sciences University of Ankara are archived in the Library of Ankara University (Volume 1-63) and Dergipark immediately following publication with no embargo.

**Editor in Chief**

Commun. Fac. Sci. Univ. Ank. Ser. A2-A3.

Ankara University, Faculty of Sciences

06100 Beşevler, ANKARA – TÜRKİYE

# C O M M U N I C A T I O N S

FACULTY OF SCIENCES  
UNIVERSITY OF ANKARA

DE LA FACULTE DES SCIENCES  
DE L'UNIVERSITE D'ANKARA

Series A2-A3: Physical Sciences and Engineering

**Volume: 66**

**Number: 2**

**Year: 2024**

---

## Research Articles

---

M.E. YEŞİLYURT, M.S. GÜZEL and E. AKÇAPINAR SEZER, A comparison of machine learning methods for queue length detection .....	132
T. ASUROGLU, Enhancing precision in proton therapy: utilizing machine learning for predicting Bragg curve peak location in cancer treatment.....	140
A. YONAR and Ö. ÖZALTIN, Automatic recognition of coffee bean varieties based on pre-trained CNN architectures .....	162
K. AÇICI, Hemoglobin value prediction with bayesian optimization assisted machine learning models.....	176
I. NAVRUZ, C. DILSIZ, E.S. ORTAK, S.N. BOYRAZ, Optical fiber bending sensor based on speckle pattern imaging .....	201
V. YILMAZ, Sending pictures over radio systems of the trail cam in border security and directing uavs to the right areas .....	214
Y.O. GÜNAYDIN, M. SAHIN and L. AYDIN, Search potential of the high energy-large hadron collider for spin-1/2 excited quarks in di-jet final state.....	228