Konya Mühendislik Bilimleri Dergisi., c. 13, s. 4, 991-998, 2025
Konya Journal of Engineering Sciences v. 13, n. 4, 991-998, 2025
ISSN: 2667-8055 (Electronic)
DOI: 10.36306/konjes.1662027

EXTRACTING A MUCH WIDER ABSORPTION SPECTRAL RANGE FROM A NARROWER RANGE USING TRANSFORMER MODEL

[1,*] **Enes EKEN**, [2] **İsmail BAYRAKLI**

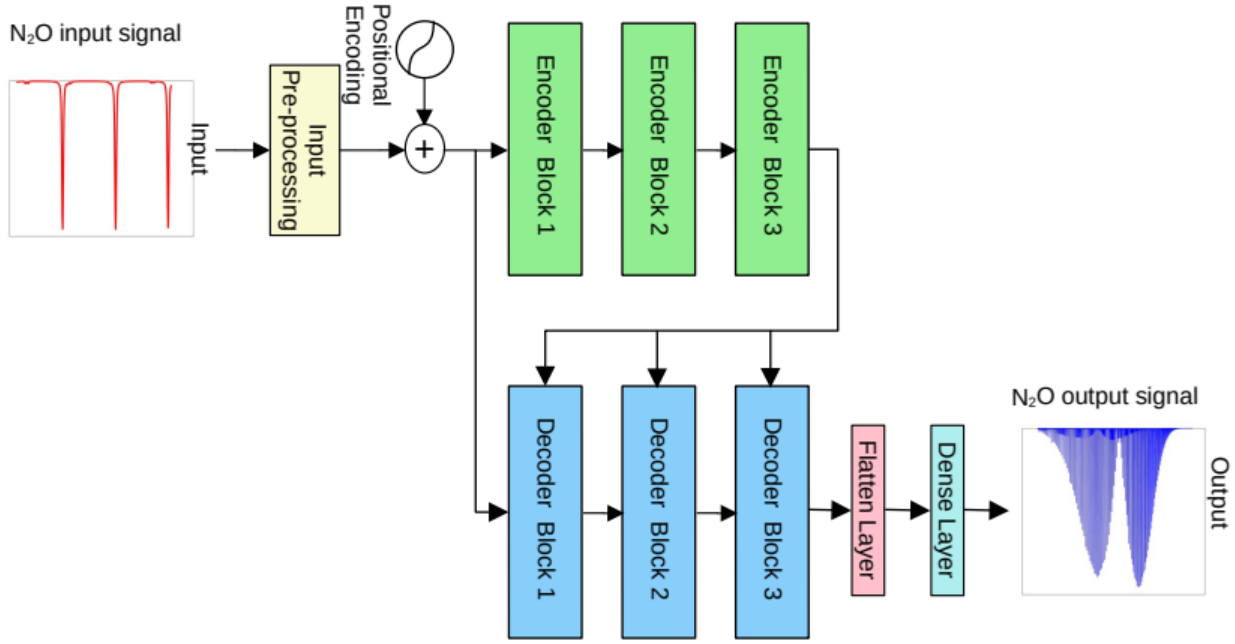
[1, 2] *Aksaray University, Engineering Faculty, Electrical and Electronics Engineering Department, Aksaray, TÜRKİYE*

[1] eneseken@aksaray.edu.tr, [2] ismailbayrakli@aksaray.edu.tr

Highlights

- A novel sensor system based on Transformer
- Rapid acquisition of a very wide spectral range at high resolution
- Transformer based model for obtaining wide absorption spectral range

Graphical Abstract



The overall architecture of the proposed model consists of three identical encoder and decoder blocks

***Corresponding Author:** Enes EKEN, eneseken@aksaray.edu.tr

Konya Mühendislik Bilimleri Dergisi., c. 13, s. 4, 991-998, 2025
Konya Journal of Engineering Sciences v. 13, n. 4, 991-998, 2025
ISSN: 2667-8055 (Electronic)
DOI: 10.36306/konjes.1662027

EXTRACTING A MUCH WIDER ABSORPTION SPECTRAL RANGE FROM A NARROWER RANGE USING TRANSFORMER MODEL

[1,*] Enes EKEN, [2] İsmail BAYRAKLI

[1,2] *Aksaray University, Engineering Faculty, Electrical and Electronics Engineering Department, Aksaray, TÜRKİYE*

[1] eneseken@aksaray.edu.tr, [2] ismailbayrakli@aksaray.edu.tr

(Received: 20.03.2025; Accepted in Revised Form: 08.07.2025)

ABSTRACT: Obtaining a wide absorption spectral range rapidly at high resolution plays a crucial role in scientific research and practical applications. The main motivation of this work is to investigate the rapid acquisition of a very wide spectral range at high resolution by using a much narrower absorption spectral range. To accomplish this, we proposed a novel transformer-based approach that produces a wider absorption spectral range (100 cm^{-1}) with high resolution in milliseconds from a narrow absorption spectral range (2 cm^{-1}). A distributed feedback quantum cascade laser (DFB QCL) together with a multi-pass cell was used to obtain the absorption lines. To evaluate the performance of the sensor, N_2O was selected as the target molecule.

Keywords: *Machine Learning, Deep Learning, Transformer, Lasers, Absorption*

1. INTRODUCTION

To analyze the molecule to be examined in laser spectroscopy methods, the absorption lines of this molecule must be revealed. For this purpose, wavelength-tunable lasers, such as external-cavity (EC) [1] or distributed feedback (DFB) [2] quantum cascade lasers (QCLs), can be used. In such applications, rapid scanning of a wide spectral area is desirable. In an EC-QCL setup, the emission spectral range of the used QCL can be fully scanned by rotating the diffraction grating with a motorized actuator. A broadly coarse wavelength tuning range can reach up to 100 cm^{-1} [1], depending on the gain curve of the QCL chip used. However, due to the mechanical motion of the motor, the total scanning time reaches a few minutes. The resolution of the absorption spectra of molecules depends on the speed of the motor. To achieve a better resolution, the speed of the motor has to be reduced. In this case, scanning time may take longer. To achieve a high scanning speed in the millisecond range, a piezo element mounted on the diffraction grating can be used, keeping the motorized rotation stage stationary. However, in this case, a fine-tuning range of approximately 5 cm^{-1} can be reached [3]. A similar spectral tuning range can be achieved in milliseconds by controlling the current or temperature of the DFB-QCL chip [2]. In the EC-QCL setup with intracavity acoustic-optic frequency shifter, wavelength could be tuned rapidly in a broad tuning bandwidth of 33 cm^{-1} with in 50 μs by varying the acoustic frequency. However, due to the broader line width in this setup, high-resolution gas spectroscopy is not possible [4]. In order to address this issue, we proposed a novel supervised machine learning approach by utilizing the Transformer[11].

In the past decade, we have witnessed the rapid rise of supervised machine learning methods as a solution to many complex problems in different fields spanning from image classification [5] to object detection [6], from credit card fraud detection [7] to molecule concentration prediction leveraging the absorption spectra [8]. The distinguishing feature of a supervised machine learning is that it requires a paired input-output dataset in order to train the model. For instance in image classification problem these input-output pairs are the images and their corresponding labels. Similarly for molecule concentration prediction, absorption spectra will be the input signal, and the output is the concentration value of the related molecule. The deep learning model should be designed carefully for the specific task considering the inductive biases of the used layers (e.g., CNN [9], LSTM [10], or Transformer [11]). Afterward, an objective function which measures the performance of the model should be determined. While this

***Corresponding Author:** Enes EKEN, eneseken@aksaray.edu.tr

function can be categorical cross entropy for image classification, if the task is real valued prediction, it can be mean squared error. Later on, the model should be trained by optimizing the objective function over this train dataset. As an optimization method, variants of stochastic gradient descent such as Adam [12], Rmsprob [13] can be used.

In the present study, a novel Transformer [11] based approach that produces a wider absorption spectral range (100 cm^{-1}) with high resolution in milliseconds from a narrow absorption spectral range (2 cm^{-1}) is reported. The narrow absorption spectral range is produced using a DFB-QCL and a multi-pass absorption cell. When this spectral area is given as input to our model, a wider spectral range is obtained. The most important advantage of this method is that it allows reaching a very wide spectral area very quickly (ms), without requiring any mechanical movement. In this way, large molecules (molecules consisting of more than four atoms) can be easily detected and quantified. So far, such a wide spectral area has not been achieved in such a fast time using EC-QCL. To the best of our knowledge, this is the first time this achievement is obtained.

2. PRELIMINARIES

In sequence to sequence and transduction problems where the inputs and outputs are time dependent signals e.g., laser absorption spectra or words as in natural language processing, Recurrent Neural Networks (RNN) [14] and particularly its popular variants Long short-term memory (LSTM) [15] and Gated Recurrent Neural Networks [16] have been commonly used as defacto choices for long years, thanks to their past dependent memory capacity. RNNs process the input data sequentially, and are trained by using well known back-propagation through time algorithm. Although these models show successes in their era, they have certain drawbacks. Vanishing and exploding gradients [15] are two of the biggest issues that RNNs have to deal along with sequential computation costs. Due to the sequential nature of RNNs, they do not fully benefit Graphics Processing Units (GPUs). In order to relief the sequential computation cost and draw global dependencies, Vaswani et al. [11] proposed Transformer model. Although Transformer was initially proposed in Natural Language Processing domain for addressing the one of the most challenging tasks such as neural machine translation by leveraging its attention mechanism, its success made it a prominent candidate in other domains and adopted in computer vision, image classification [17], object detection [18], and gas concentration prediction using absorption spectra [8].

The overall architecture of the Transformer [11] can be seen in Figure 1. It adopts the general encoder-decoder structure through an attention mechanism but does not rely on any recurrence or convolution. An encoder block consists of two sub-layers, namely Multi-Head Attention and position wise fully connected feed-forward network. Feed-forward network comprises of two linear layers and performs ReLU activation function over its inputs. For a stable learning, each sub-layer is followed by a normalization layer [19]. In addition, each sub-layer has a residual connection [20] around itself.

Transformer decoder is also similar to encoder except the third multi-head attention sub-layer which receives the input from encoder output and previous sub-layer of the decoder. Layer normalization and residual connection are also implemented in the decoder.

Since the transformer does not have any recurrence, conceptually it is order agnostic. Therefore a positional encoding is injected to input embedding which is defined as;

$$PE_{POS,2i} = \sin\left(\frac{pos}{10000^{2i/d_m}}\right) \qquad (1)$$

$$PE_{POS,2i+1} = \cos\left(\frac{pos}{10000^{2i/d_m}}\right)$$

For a sequence of length n and embedded dimension of $\mathbf{d_m}$, assume $\mathbf{X} \in R^{n \times dm}$ is sum of embedding of the sequence and positional encoding of it. Then **Q**uery , **K**ey, and **V**alue will be the linear transformation of $\mathbf{X}$ such that $\mathbf{Q} = \mathbf{XW^q}$, $\mathbf{K} = \mathbf{XW^k}$, $\mathbf{V} = \mathbf{XW^v}$, where $\mathbf{W^q} \in R^{dm \times dk}$, $\mathbf{W^k} \in R^{dm \times dk}$, and $\mathbf{W^v} \in R^{dm \times dv}$. As attention

mechanism, scaled dot-product attention was preferred in the transformer. For given **Q**, **K**, and **V**, attention can be calculated as;

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{(d_k)}}\right)V \tag{2}$$

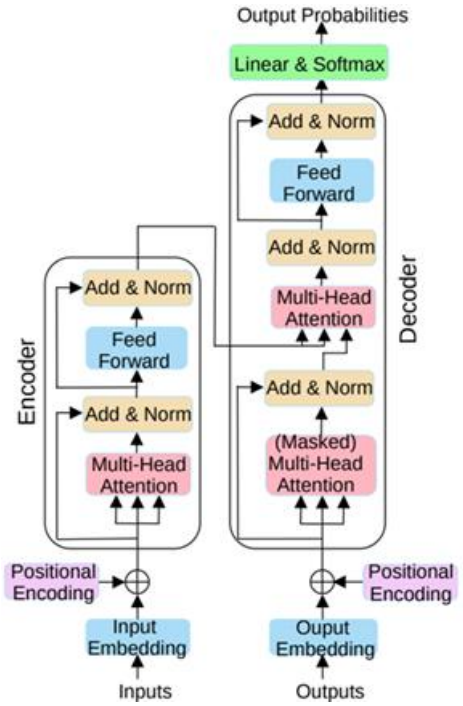


Figure 1. Overview of the Transformer architecture which consists of encoder and decoder blocks through scaled dot product attention mechanism

In order to increase the computational capacity of the model, Vaswani et al. [11] proposed to use multi-head attention instead of the single attention layer, since with this way extracting multiple correlation from the input data will be possible. To do so, with distinct matrices $\mathbf{W}_i^q$, $\mathbf{W}_i^k$, and $\mathbf{W}_i^v$ multi-head attention is defined as;

$$MultiHead(Q, K, V) = Concat(head_1, \ldots, head_h)W^O, \tag{3}$$

where,

$$head_i = Attention(QW_i^q, KW_i^k, VW_i^v), \tag{4}$$

and $\mathbf{W}^o \in R^{hdv \times dm}$.

The second multi-head attention layer of the transformer decoder receives the **K** and **V** from the last encoder output and **Q** from the previous Add & Norm layer of the decoder.

3. EXPERIMENTAL SETUP

The established setup is shown in the figure 2. It has been discussed in detail in our newly published article [21]. The experimental setup can be briefly described as follows. A continuous-wave (CW) DFB QCL (Thorlabs, QD4474CM1AS) emitting at 4474 nm wavelength was used as an light source to target the absorption lines of N_2O. A diode/TEC controller (Thorlabs, ITC4002 - QCL) was employed to drive the

laser. The laser beam collimated by an aspheric lens (Thorlabs, C093TME-E) was sent into a compact multi-pass absorption cell (Thorlabs, MGC3C/M-P01) with a 3.2 m optical path length, to analyze the molecules under investigation. The laser beam emerging from the cell was measured by a detector (Thorlabs, PDA07P2). The electrical signal generated by the detector was sent to the PC via DAQ.

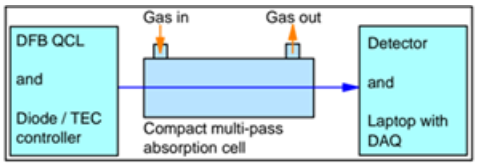


Figure 2. Schematic of the sensor. DFB QCL: Distributed Feedback Quantum Cascade Laser, TEC: Thermoelectric cooler, DAQ: Data acquisition card

In this work, in order to show the effectiveness of the proposed approach, we chose N_2O as the target molecule. While the input wavenumber of the molecule is between 2235.6 - 2237.7 cm^{-1} , the output wavenumber will be in the range of 2170 - 2272 cm^{-1} , which is approximately 48× wider as depicted in Figure 3. The model we proposed for this task is presented in Figure 4, which utilizes the Transformer as backbone. Specifically, three identical encoder blocks are stacked together. Each encoder block has 2 multi-head attention layers and embedded dimension of feed forward network is set to 128. ReLU is used for activation function. Similarly we used three identical decoder blocks sequentially connected. The embedded dimensions, number of heads and activation functions are the same with the encoder parts. In this setting, an encoder has 1.2 million training parameters, while a decoder has 2.3 million parameters. The output of the last decoder is flattened and connected to a dense layer with 25502 units which represents the dimension of N_2O output signal. The mean squared error loss function is optimized with rmsprop [13] optimizer for 100 epochs. A dynamic learning rate scheduler which starts with $1e-5$ and gradually reduces until $1e-7$ is preferred since it gave a better performance. The model is trained with Tensorflow framework using four RTX-A6000 GPUs. The validation and training loss curves are depicted at Figure 5.

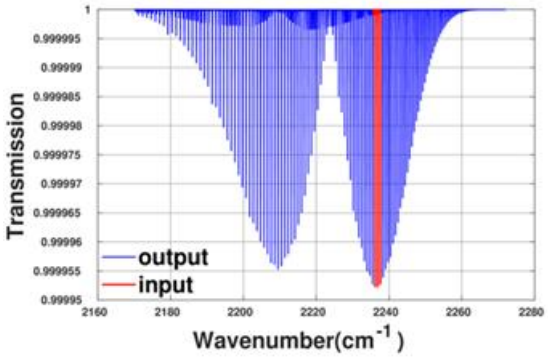


Figure 3. Input and output transmission signals of the propose model. As input we only use a narrow signal in order to predict wider version of it.

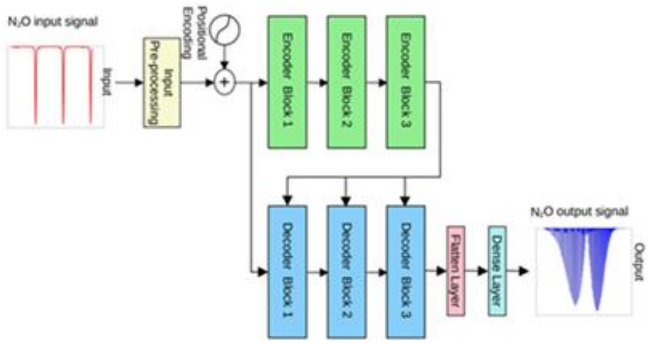


Figure 4. The overall architecture of the proposed model consists of three identical encoder and decoder blocks. While the absorption spectral range of the input signal is narrow, its corresponding output is 48X wider.

4. RESULTS AND DISCUSSION

In order to prepare the dataset, using HITRAN database, we created 10.000 input N_2O signals for varying concentrations ranging between 0-2 ppm at 2235.6 - 2237.7 cm^{-1} wavenumber. For the same concentration values we also created 10.000 output N_2O signals at 2170 – 2272 cm^{-1} wavenumber. While 80 % of the dataset is chosen for training the model, the rest of the dataset is equally partitioned for validation and the test set. There are many design space parameters which should be decided before training the model known as hyper-parameters such as the number of encoder and decoder blocks, embedded dimensions, number of multi-head attention layers, learning rates so on forth. In order to choose the best hyper-parameters, we applied K-fold cross-validation.

In practical applications, there will be many internal and external noise sources i.e., instrument based noises, parasitic reflections, laser intensity oscillations just a few of them. These noises will dramatically the change shape of the spectra, and therefore they must be taken into account. Based on the Central Limit Theorem (CLT), sum of the independent random variables (or sum of the possible parasitic effects including the ones aforementioned above) approaches the Cumulative Distribution Function (CDF) of a normal (Gaussian) random variable. Hence, the total effect of any possible noises can be safely modeled as Gaussian noise. For this reason, we added Gaussian noises with zero mean and varying standard deviations to the original input signals. Following that, we applied normalization and Short-time Fourier transformation (STFT) to the noisy input signal as presented at [22].

Figure 5. The proposed model was trained for 100 epochs for optimizing the mean squared error loss function with rmsprop. A dynamic learning rate scheduler is used. After 100 epochs the loss curves reached saturation. The model is trained with Tensorflow framework using four RTX-A6000 GPUs.

The test case of the proposed model is presented in Figure 6. For a randomly chosen concentration value up to 2 ppm, the input signal of N_2O as depicted at Figure 6-a, is given to our model, which subsequently produces an output signal as shown in Figure 6-b. For the sake of comparison, we added the theoretical target signal in the same figure, which exactly matches the prediction output.

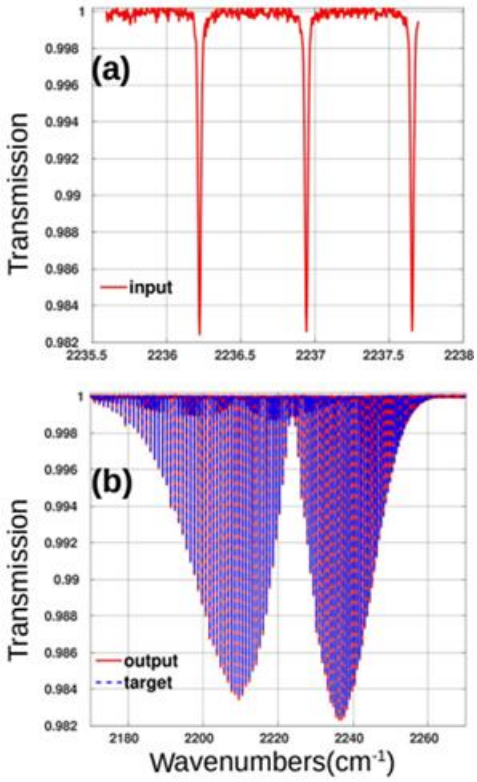


Figure 6. The model receives the narrow input signal(a), and produces wider output signal(b). The output signal produced by the proposed model is approximately 48X wider than the input signal. For the comparison purpose, we also depicted the theoretical target signal in the figure. As can be seen from (b) , the target and predicted output signals match excatly.

5. CONCLUSION

To sum up, a novel sensor system based on the Transformer approach has been demonstrated to generate a wider absorption spectral range (100 cm−1) with high resolution in milliseconds from a narrow absorption spectral range (2 cm−1). Our transformer based model produces 48X wider absorption spectral range as output from a narrower absorption spectal range. The performance of the sensor has been evaluated using N_2O molecules. A DFB-QCL combined with a compact multi-pass cell has been used to target the absorption lines of N_2O. In future studies, the used method will be extended to the rapid and easy analysis of large molecules containing more than four atoms. The results obtained in this study can be used for other molecules. In this case, the laser to be used should be selected accordingly.

Declaration of Ethical Standards

Authors followed all ethical guidelines including authorship, citation, data reporting, and publishing original research.

Credit Authorship Contribution Statement

Author1, as first author, is responsible of machine learning related parts including model design, training, and testing in addition to writing of manuscript. Author2 is responsible of laser sensor system in addition to writing laser related things.

Declaration of Competing Interest

The authors declare that there is no conflict of interest.

Funding / Acknowledgments

This work was supported by the Scientific and Technological Research Council of Turkey (project number: 124E058).

Data Availability

All data generated or analyzed during this study are included in this published article.

REFERENCES

[1] Y. Matsuoka, et al., "External-cavity quantum cascade laser using intra-cavity out-coupling," Optics Letters 43, 3726-3729 (2018).
[2] I. Bayrakli, "A portable N_2O sensor based on quartz-enhanced photoacoustic spectroscopy with adistributed-feedback quantum cascade laser for medical and atmospheric applications," Optical and Quantum Electronics 53, 642 (2021).
[3] I. Bayrakli, and H. Akman, "Ultrasensitive, real-time analysis of biomarkers in breath using tunable external cavity laser and off-axis cavity-enhanced absorption spectroscopy," Journal of Biomedical Optics 20, 037001 (2015).
[4] I. Bayrakli, "External cavity quantum cascade lasers without anti-reflection coating with intracavity and extracavityacoustic-optic frequency shifter for fast standoff detection," Optics and Laser Technology 148, 107747 (2022).
[5] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).
[6] J. Redmon, "You only look once: Unified, real-time object detection," In Proceedings of the IEEE conference on computer vision and pattern recognition (2016).
[7] F. Hosein, and H. Abbasimehr, "A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection," Expert Systems with Applications 217, 119562 (2023).
[8] I. Bayrakli, and E. Eken, "Compact laser spectroscopy-based sensor using a transformer-based model for analysis of multiple molecules," Applied Optics 63, 6941-6947 (2024).
[9] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE 86, 2278-2324 (1998).
[10] S. Hochreiter, "Long Short-term Memory," Neural Computation MIT-Press (1997).
[11] A. Vaswani, "Attention is all you need," Advances in Neural Information Processing Systems (2017).
[12] D.P. Kingma, and B. Jimmy "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).
[13] T. Tieleman, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA:Neural networks for machine learning," 4, 2 (2012).
[14] J. L. Elman, "Finding structure in time. Cognitive science," 14, 2 (1990).
[15] S. Hochreiter, "Long Short-term Memory," Neural Computation MIT-Press (1997).
[16] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio. "Empirical evaluation of gated recurrent neural

networks on sequence modeling," arXiv preprint arXiv:1412.3555 (2014).
[17] L. Ze, et al., "Swin transformer: Hierarchical vision transformer using shifted windows," In Proceedings of the IEEE/CVF international conference on computer vision, 10012-10022. (2021).
[18] C. Nicolas, et al., "End-to-end object detection with transformers," In European conference on computer vision," Cham: Springer International Publishing, 213-229 (2020).
[19] J. L. Ba, "Layer normalization," arXiv preprint arXiv:1607.06450 (2016).
[20] H. Kaiming, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778 (2016).
[21] Bayrakli, and E. Eken, "Compact laser spectroscopy-based sensor using a transformer-based model for analysis of multiple molecules," Applied Optics, 63, 6941-6947 (2024).
[22] I. Bayrakli, and E. Eken. "A novel breath molecule sensing system based on deep neural network employing multiple-line direct absorption spectroscopy," Optics and Laser Technology 158 108918 (2023).