

e-ISSN: 2980-177X

JCEES

Volume: 2 Issue: 2 Year: 2024

**Journal of
Computer &
Electrical and
Electronics
Engineering
Sciences**

Journal of

Computer & Electrical and Electronics Engineering Sciences

EDITORIAL BOARD

EDITOR-IN-CHIEF

Assist. Prof. Fuat TRK

Department of Computer Engineering, Faculty of Architecture and Engineering, Kırıkkale University, Kırıkkale, Türkiye

ASSOCIATE EDITORS-IN-CHIEF

Assist. Prof. Hseyin AYDİLEK

Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Kırıkkale University, Kırıkkale, Türkiye

Dr. Mahmut KILIÇASLAN

Department of Computer Technologies, Vocational School of Nallıhan, Ankara University, Ankara, Türkiye

Assoc. Prof. Selim BUYRUKOĐLU

Department of Computer Engineering, Faculty of Engineering, Çankırı Karatekin University, Çankırı, Türkiye

EDITORIAL BOARD

Dr. Alper Nuri METİN

Department of Electronics and Communications, Vocational School of Kırıkkale, Kırıkkale University, Kırıkkale, Türkiye

Prof. Aydın ÇETİN

Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye

Assist. Prof. Ayhan AKBAŞ

Department of Computer Engineering, Faculty of Engineering, Abdullah Gl University, Kayseri, Türkiye

Assist. Prof. Ebru AYDOĐAN DUMAN

Department of Computer Information Systems Engineering, Faculty of Bucak Technology, Burdur Mehmet Akif University, Burdur, Türkiye

Assist. Prof. Elvan DUMAN

Department of Software Engineering, Faculty of Bucak Technology, Burdur Mehmet Akif University, Burdur, Türkiye

Assist. Prof. Enes AYAN

Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Kırıkkale University, Kırıkkale, Türkiye

Assoc. Prof. Erinç KARATAŞ

Department of Computer Technologies, Vocational School of Elmadağ, Ankara University, Ankara, Türkiye

Assist. Prof. Faruk ULAMIŞ

Department of Electronics and Automation, Vocational School of Hacılar Hseyin Aytemiz, Kırıkkale University, Kırıkkale, Türkiye

Assoc. Prof. Fatih KORKMAZ

Department of Electrical and Electronics Engineering, Faculty of Engineering, Çankırı Karatekin University, Çankırı, Türkiye

Assoc. Prof. Hseyin POLAT

Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye

Assoc. Prof. İbrahim Alper DOĐRU

Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye

Journal of

Computer & Electrical and Electronics Engineering Sciences

EDITORIAL BOARD

Assist. Prof. İsmail ATACAK

Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye

Prof. İsmail Rakıp KARAS

Department of Computer Engineering, Faculty of Engineering, Karabük University, Karabük, Türkiye

Assoc. Prof. Levent GÖKREM

Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Tokat Gaziosmanpaşa University, Tokat, Türkiye

Assist. Prof. Mehmet GÜÇYETMEZ

Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences, Sivas University of Science and Technology, Sivas, Türkiye

Assoc. Prof. Muhammet Tahir GÜNEŞER

Department of Electrical and Electronics Engineering, Faculty of Engineering, Karabük University, Karabük, Türkiye

Assoc. Prof. Murat LÜY

Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Kırıkkale University, Kırıkkale, Türkiye

Assist. Prof. Mürsel Ozan İNCETAŞ

Department of Computer Technologies, Vocational School of Alanya Ticaret ve Sanayi Odası, Alanya Alaaddin Keykubat University, Antalya, Türkiye

Assist. Prof. Mustafa TEKE

Department of Electrical and Electronics Engineering, Faculty of Engineering, Çankırı Karatekin University, Çankırı, Türkiye

Assist. Prof. Mustafa KARHAN

Department of Computer Engineering, Faculty of Engineering, Çankırı Karatekin University, Çankırı, Türkiye

Assist. Prof. Mustafa Yasin ERTEN

Department of Electrical and Electronics Engineering, Kırıkkale University, Kırıkkale, Türkiye

Prof. Necaattin BARIŞCI

Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye

Prof. Necmi Serkan TEZEL

Department of Electrical and Electronics Engineering, Faculty of Engineering, Karabük University, Karabük, Türkiye

Assoc. Prof. Ramazan Kürşat ÇEÇEN

Department of Aircraft Technology, Vocational School of Eskişehir, Eskişehir Osmangazi University, Eskişehir, Türkiye

Assist. Prof. Rukiye KARAKIŞ

Department of Software Engineering, Faculty of Technology, Cumhuriyet University, Sivas, Türkiye

Assist. Prof. Saadin OYUCU

Department of Computer Engineering, Faculty of Engineering, Adıyaman University, Adıyaman, Türkiye

Dr. Salih DEMİR

Department of Computer Engineering, Faculty of Open and Distance Education, Ankara University, Ankara, Türkiye

Assist. Prof. Sevilay VURAL

Department of Internal Medicine Sciences, School of Medicine, Yozgat Bozok University, Yozgat, Türkiye

Assoc. Prof. Sinan TOKLU

Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye

Computer & Electrical and Electronics Engineering Sciences

EDITORIAL BOARD

Dr. Ufuk TANYERİ

Department of Computer Technologies, Vocational School of Nallıhan, Ankara University, Ankara, Türkiye

Dr. Yunus KÖKVER

Department of Computer Technologies, Elmadağ Vocational School, Ankara University, Ankara, Türkiye

Assist. Prof. Zafer CİVELEK

Department of Electrical and Electronics Engineering, Faculty of Engineering, Çankırı Karatekin University, Çankırı, Türkiye

ENGLISH LANGUAGE EDITOR

Assoc. Prof. Esra Güzel TANOĞLU

Department of Molecular Biology and Genetics, Institute of Health Sciences, University of Health Sciences, İstanbul, Türkiye

STATISTICS EDITOR

Assoc. Prof. Turgut KÜLTÜR

Department of Physical Therapy and Rehabilitation, Faculty of Medicine, Kırıkkale University, Kırıkkale, Türkiye

DESIGN

Şerife KUTLU

E-mail: mha.editoroffice@gmail.com

Journal of

Computer & Electrical and Electronics Engineering Sciences

CONTENTS

Volume: 2 Issue: 2 Year: 2024

ORIGINAL ARTICLES

Sentiment analysis with machine learning for drug reviews	35-45
<i>Bozkurt, M.O., Yaman, Y., & Horasan, F.</i>	
Steering angle prediction in autonomous vehicles: a deep learning approach combining VGG16 and LSTM.....	46-51
<i>Karadeniz, A.M., & Koçak, N.F.</i>	
Determination of I-V curves for photovoltaic generator by an analytical method: Akbaba Model.....	52-55
<i>Gökşenli, N., Bektaş, E., & Tümay, M.</i>	
Wireless power transfer systems and wireless charging design between electric vehicles.....	56-61
<i>Bilgiç, E., Bilgiç, H., & Kantaroğlu, E.</i>	
Vision assistant for visually impaired individuals.....	62-66
<i>Yıldırım, İ., Demir, İ., Peker, İ., Yılmaz, M.E., Güneyli, O., Aktaş, E.Ş., Özer, E., & Kantaroğlu, E.</i>	

Sentiment analysis with machine learning for drug reviews

✉ Muhammed Oğuzhan Bozkurt¹, ✉ Yağız Yaman², ✉ Fahrettin Horasan²

¹Department of Software Engineering, Faculty of Engineering, İstanbul Gedik University, İstanbul, Türkiye

²Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Kırıkkale University, Kırıkkale, Türkiye

Cite this article: Bozkurt, M.O., Yaman, Y., & Horasan, F. (2024) Sentiment analysis with machine learning for drug reviews. *J Comp Electr Electron Eng Sci*, 2(2), 35-45.

Corresponding Author: Fahrettin Horasan, fhorasan@kku.edu.tr

Received: 29/08/2024

Accepted: 22/09/2024

Published: 29/10/2024

ABSTRACT

In the treatment of the diseases, the fact that individuals use drugs independently from doctors without appropriate consultation causes their health status to become worse than normal. This article aims to conduct a sentiment analysis over the comments of individuals about the drug in case they use drugs without consultation. Within the scope of this study, patients' comments about drugs were vectorized using Bow and TF-IDF algorithms, sentiment analysis was made, and the predicted sentiments were; it was evaluated with precision, recall, f1score, accuracy and AUC score. As a result of the evaluations, the most successful result was obtained in the TF-IDF method. This result is the result of the linear support vector classifier algorithm with an accuracy value of 93%.

Keywords: Drug, sentiment analysis, machine learning, natural language processing, smote

INTRODUCTION

One of the worldwide problems with the pandemic is the insufficient number of specialist doctors. This situation sometimes delays patients' access to appropriate diagnosis and treatment. It takes between 6 and 12 years for an average doctor to acquire the necessary qualifications, so the number of qualified doctors cannot be increased rapidly in a short time (Garg, 2021). It should be considered that telemedicine applications can accelerate patients' access to appropriate treatment in global health crises such as pandemics (Dinakaran et al., 2020).

In addition, clinical errors are quite common nowadays. According to the findings obtained from the studies, 200 thousand people in China and more than 100 thousand people in the USA are exposed to wrong treatment due to misdiagnosis and prescriptions (Garg, 2021). On the other hand, in the medical world, experts make more than 40% errors while writing prescriptions, mostly due to the limited knowledge of specialists in the treatment of the disease (Wittich et al., 2014). In addition, new drugs emerging every day led to the emergence of tests and new studies for clinical staff. This situation makes it increasingly difficult for doctors to determine the appropriate treatment and drug for the patient, depending on the indication and clinical history.

The rapid development of the web-based business industry has made product reviews an indispensable factor in purchasing. Individuals browse product reviews and alternative websites when deciding on the product to buy. Much of the past

research has focused on ratings and recommendations on e-commerce sites and has rarely been studied in the field of medical care or clinical treatments. Recently, the use of online diagnostic websites has increased considerably. According to a survey conducted by the Pew American Research Center in 2013, approximately 60% of individuals have searched online for health-related topics this year, and approximately 35% of these individuals have received help from the web in diagnosing their disease (Fox et al., 2013). Considering these situations, it is vital to develop a drug recommendation system to assist experts and ensure that individuals reach the appropriate medicine for their disease.

Recommendation systems are systems designed to facilitate and accelerate users' access to the products they need (Jalili et al., 2018). In the drug recommendation system, drug recommendations are made to users using sentiment analysis and feature engineering. While sentiment analysis is the extraction of some emotional data such as opinions and attitudes from the relevant text (Kaur et al., 2017); feature engineering is optimizing existing data and extracting more features (Oyamada, 2019). Considering all these evaluations, it should not be overlooked that emotion analysis should be performed in the most accurate way before the recommendation system to be developed for patients to reach the appropriate treatment. This study aims to carry out the sentiment analysis to be performed with the highest possible accuracy.

In the continuation of the study, there is a literature review section in which the researches related to the subject are examined, the methodology section in which the methods applied within the scope of the study are explained and analyzed, the results section where the results of the applied methods are examined, and the discussion and solution sections in which the study is examined.

In the continuation of the study, there is a literature review section where the researches related to the subject are examined, the methodology section where the methods applied within the scope of the study are explained and examined, the results section where the results of the applied methods are examined, and the discussion and solution sections where the study is examined.

Literature Review

With the increase in the development of artificial intelligence, machine learning and deep learning methods have started to be applied in recommendation systems. Recommendation systems today it is frequently used in the travel industry, e-commerce sites, restaurant applications and TV series-film sites. However, in the field of drug recommendation medical expressions such as drug reviews, disease names, reactions, and synthetic names are complex and difficult to understand, leading to limited studies involving sentiment analysis in this area (Tekade et al., 2016).

In this article, in which a recommendation system that applies sentiment analysis technologies in drug reviews was created, a decision support platform was designed to assist patients in their drug selection. First, a rating of drugs was created by performing emotion analysis with drug reviews. Second, how useful drug reviews are for users, the patient's conditions, and the glossary sensitivity polarity of drug reviews are taken into account. These factors were then included in the recommendation system to list suitable drugs. Finally, hyper parameter optimization is performed for each algorithm to achieve higher performance (Hossain et al., 2020).

In this study, in which a recommendation system was developed using product images, Amazon Apparel database with clothing data was used. NLP technologies and CNN are used to predict similar products. The title of the product was chosen as the main attribute for NLP analysis and product recommendation. CNN is used to generate a feature vector from product images, and all other vectors are used to estimate this vector. By calculating the distances between the vectors of all products, the products with the least distance are recommended. VGG-16 architecture is used to extract features from images (Sharma et al., 2021).

In another study, it was assumed that the recommended drug should be determined according to the patient's immunity. For example, if the patient's immunity is low, reliable drugs should be recommended at this point. The main purpose of the study is to protect a patient from infection by making use of the patient's clinical information. Risky situations such as the patient's weaknesses and allergies are evaluated and scored according to the effect of these conditions. The system calculates the risk level after comparing the risk factors, allowing doctors to easily make inferences about the patient's condition. As a result of the study, a web-based prototype system was also created that uses a decision support system that helps doctors choose primary care drugs. However, as

in the previous study, it should be noted that this study did not design a recommendation system based on sentiment analysis and machine learning (Shimada et al., 2005).

In this article, in which vaccine hesitancy is examined and a solution to this problem is tried to be produced by sentiment analysis, articles related to different vaccine hesitations made in the last 11 years have been examined and analyzed. It is intended to develop the application of sentiment analysis on the most important literature findings (Alamoodi et al., 2021).

One of these studies presents GalenOWL, an online framework based on semantically augmented semantic web technologies to help professionals discover details about drugs. GalenOWL describes a framework that recommends medication for a patient based on the patient's infection, sensitivities, and drug interactions. However, this work requires advanced medical knowledge and many complex and memory-intensive queries. The recommendations generated in the system are based on complex queries that do not rely on sentiment analysis and machine learning, but mostly consider relationships between data (Doulaverakis et al., 2012).

In this study, in which the movie recommendation system developed using sentiment analysis is presented, cosine similarity is used for the developed recommendation system. The main purpose of the study is to present the content they need to end users through semi-structured data on the internet. As a result of the study, it was concluded that cosine similarity provides better and more efficient results for a recommendation system (Khatter et al., 2021).

In another study focusing on movie recommendations, sentiment analysis was performed using the LDA method, using user comments. The sentiment analysis has been used to develop a content-based recommendation system. In this study, BERT technique was used to train emotion classification models and LDA technique was used for modeling subjects (Zhang et al., 2022).

Zhang et al. (2014) has developed a cloud-assisted drug recommendation system (CADRE). According to patients' side effects, CADRE recommends the most relevant prescription drugs. This proposed framework was initially built on collaborative filtering techniques as indicated by the functional identification data of drugs. The model is then transformed into a cloud-powered approach that uses tensor decomposition to improve the quality of the drug recommendation experience.

In this study, in which a tourism recommendation system was developed, sentiment analysis of user comments on TripAdvisor was used to identify touristic places that might be of interest to tourists. After the comments are processed, they are clustered semantically and sentiment analysis is performed. These comments are then used to extract the attractions of the attractions. Finally, the user is recommended the most suitable tourist attractions. In addition, the system uses some contextual information such as time, location and weather to filter items and improve the quality of recommendations (Abbasi-Moud et al., 2021).

In another study, a universal drug recommendation system framework was designed and implemented, applying data mining technologies to the recommendation system. Drug

recommendation system consists of database system module, data preparation module, recommendation model module, model evaluation and data visualization module. Based on the diagnosis data, support vector machine (SVM), BP neural network algorithm and ID3 decision tree algorithm were used for drug recommendation system. Although data mining methods were used in the study, sentiment analysis based on user comments was not performed (Bao et al., 2016).

In this study, in which a different approach to recommendation systems is presented, a hierarchical approach is presented to improve the performance of e-commerce recommendation systems. This approach, called DeepIDRS, has a two-level hierarchical structure. At the first level, a bidirectional encoder is used to represent the textual information of an item efficiently and accurately. At the second level, an attention-based recommendation system is used, which uses the item representations produced at the first level. In addition, the developed DeepIDRS approach has been compared with other existing approaches. As a result, it has been observed that this approach improves the performance of the recommendation system (Islek et al., 2022).

In this article, which examines the basic approaches used for sentiment analysis, important concepts related to sentiment analysis, the usage areas of sentiment analysis and the preprocessing process required for the realization of sentiment analysis are examined. In addition, models that can be used for machine learning are also examined (Kaur et al., 2017).

Pereira et al. (2020) presented a study aiming to compare these techniques by analyzing the filtering techniques commonly used in recommendation systems. In addition, criteria such as weighting, frequency, distance and similarity used in recommendation systems were also evaluated.

In another study, the reflections of people's professions on social media platforms were analyzed. Based on John Holland's theory of occupational choice, the study aims to classify occupation-related content on Twitter data of four different occupational groups. Long short-term memory (LSTM) and gated recurrent unit (GRU) models were used for classification. While the LSTM model yielded an accuracy score of 93.025%, the GRU model achieved 94.025% accuracy. These results suggest that social media posts may be consistent with individuals' occupations and that their occupations are reflected in their behavior on these platforms (Dagistanlı et al., 2023).

METHODOLOGY

In this study, the drug review dataset (Drugs.com) available in the UCI ML repository was used for sentiment analysis. The fields in this data set; name of the drug (text), condition of the drug (text), review of the patient using the drug (text), given to the drug by the patient of rating score (numerical), date of the review (date), number of users who found the review helpful (numeric) consists of features. There are a total of 215063 samples in the data set.

Within the scope of the study, a three-stage study method was adopted for sentiment analysis based on user comments. These stages are respectively; data preparation, classification, and evaluation stages.

Data Preparation

In the data preparation phase; data preprocessing methods such as checking for empty samples, checking for meaningless samples, cleaning data, extracting features from rows and categorizing data were used.

Before checking the empty samples, meaningless samples were detected and the "condition" areas of these samples were cleared. At this stage, it was determined that there were 1171 samples containing meaningless expressions. As a result of checking the empty samples, it was determined that there were a total of 2365 empty values in the "condition" field, which represents the usage status of the drug. Instead of directly deleting the rows with these empty values, care was taken to recover as much data as possible. In this way, it is envisaged to increase the performance of the sentiment analysis to be performed.

First of all, the records of 17 drugs that do not have a use case in the data set, that is, they cannot be recovered in any way, were removed from the data set. Then, 305 empty fields of the drugs with only 1 usage status were filled according to the current usage status. Finally, the use cases of the drugs were searched in the user reviews, and the empty field was tried to be filled by using the matching "condition" information in case a match was achieved. At this stage, a meticulous effort was made to fill in the blank data correctly. As a result of this process, 725 empty samples were filled. As a result, a total of 1030 data were recovered and 1318 rows containing empty samples were removed from the data set. As a result of checking the empty and meaningless samples, the number of new lines was determined as 213728. In order to improve the performance of text mining, it is extremely important to recover as much data as possible, since the "condition" field will be included in the vector to be created. In addition, the accuracy of text mining also depends on the relevance of the properties of the generated matrix to each other. For this reason, irrelevant data needs to be cleaned.

During the data cleaning phase in order to delete punctuation marks in the texts, deletion of numerical expressions, deletion of ineffective words and to obtain the roots of the words, lemmatization and stemming processes were performed. It has been observed that the vector created by lemmatization produces better results. Therefore, the study will continue with lemmatization.

Feature Extraction

After performing the data cleaning phase, feature extraction will be done using the "review" column, which contains user evaluations for sentiment analysis. The purpose of feature extraction here is to enable machine learning algorithms to work with text expressions. Each of the user ratings will be expressed as a numerical vector. This vector conversion will be done using bag of words and TF-IDF methods.

Bag of words is an algorithm used in natural language processing responsible for counting the number of all tokens in a review or document. A term or symbol can be called a word (unigram) or number of words, n-grams. A major disadvantage of the Bow model is that it creates a large matrix that is computationally costly to train (Pu et al., 2007). In this study (1.2) n-gram range was chosen.

TF-IDF is an algorithm used for statistical analyzes on text expressions in text mining and natural tooth processing. TF-IDF calculates the probability of a word being found in any text by weighting the words instead of counting them. Term frequency (TF) is the probability of the word being found in a text. Inverse document frequency (IDF) is calculated by taking the logarithm of the TF value and is inversely proportional to the TF. TF-IDF is obtained by multiplying TF and IDF values (Qaiser et al., 2018). In this study, the n-gram interval selected for TF-IDF was chosen as (1.2), just like in the bag of words algorithm.

Encoding of Data

As it is known, since many machine learning algorithms cannot learn and prediction on categorical data, these data must be encoding. Encoding of data can be done in different ways. Within the scope of this study, one-hot encoding method was preferred for encoding process. One-hot encoding provides binary representation of categorical data (Rodríguez et al., 2018).

As we mentioned before, the “condition” and “drugName” fields in the data set consist of text expressions. It is assumed that the inclusion of these fields in the machine learning models together with the “review” field will give more accurate results. After these fields were encoding with the one-hot encoding method, they were combined with the data set from which feature extraction was made through the “review” field and made ready for machine learning.

Train-Test Split

The data, created using bag of words, TF-IDF and one-hot encoding methods, is splitted into 75% training and 25% testing. The same random state value was used to obtain the same set of random numbers when splitted the data sets.

Synthetic Minority Oversampling Technique (SMOTE)

After the training-test splitting of the data, SMOTE was used to minimize the class imbalance problem. SMOTE is an oversampling technique that synthesizes existing data to balance the data. First, the difference between the examined feature vector and its nearest neighbor is taken. The result obtained is multiplied by a random number between 0 and 1. Finally, the result obtained from this multiplication is added to the examined feature vector and a synthetic new sample is formed (Feng et al., 2021).

In this study, in addition to the standard SMOTE method described above, ADASYN, SMOTE Tomek and borderline SMOTE methods were also used. Unlike SMOTE, ADASYN provides more realistic data by adding random values to synthetic samples (He et al., 2008). In the SMOTE Tomek method, Tomek links are added to the synthetic data created with SMOTE (Wang et al., 2019). In the borderline SMOTE method, the samples located on the border lines of the clusters formed by the samples belonging to the minority class type are used to create synthetic samples (Chen et al., 2021). SMOTE methods are applied only to training data. The amounts of synthetic data generated by the SMOTE methods used in the Table 1 below can be examined.

SMOTE	Class	Train (75%)	Test (25%)
No SMOTE	Negative	47991	15852
	Positive	112305	37580
	Total	160296	53432
Standart SMOTE	Negative	78613	15852
	Positive	112305	37580
	Total	190918	53432
ADASYN	Negative	78146	15852
	Positive	112305	37580
	Total	190451	53432
SMOTE Tomek	Negative	78078	15852
	Positive	111770	37580
	Total	189848	53432
Borderline SMOTE	Negative	78613	15852
	Positive	112305	37580
	Total	190918	53432

SMOTE: Synthetic minority oversampling technique

As can be seen in the Table 1, the amount of synthetic samples produced in most methods is very close to each other. In the study, the standard SMOTE method, which is the fastest method, was preferred because both time and performance costs were considered and the methods yielded very close results.

Feature Scaling

Scale differences between features can prevent machine learning from being done properly. Some machine learning algorithms that use distance-based calculations require scaling of data to work more optimally. Here, while scaling, the scale differences between the features are minimized to obtain more accurate results. In order to minimize these differences, some standardization and normalization methods are used.

In this study, scaling problem was tried to be solved by using MaxAbsScaler, RobustScaler, StandartScaler, QuantileTransformer and normalizer methods. The most appropriate scaling algorithm was determined for the machine learning models and feature extraction method used, and the method that produced the best result was used.

In MaxAbsScaler, each property is scaled according to its maximum absolute value. Each property is scaled so that its maximum absolute value is “1”. It does not shift or center the data, so that no sparsity is destroyed (Ahsan et al., 2021).

In RobustScaler, features are scaled using statistics based on outliers. When scaling, the median is removed and scales the data according to quantile intervals. The median and interquartile range can be used for scaling, as outliers will negatively affect the mean and variance (Ahsan et al., 2021).

StandardScaler is a scaling method in which the mean value of the variable distributions is converted to 0 and the standard deviation to 1, and the distribution is brought closer to the normal. It is performed by subtracting the average of the features and dividing by the standard deviation of the relevant feature. In this way, all observation units are reduced between -1 and 1.

QuantileTransformer scales features to follow a smooth or normal distribution. For a given attribute, this conversion tends to spread out the most frequent values. In addition, this

method reduces the negative effects of outliers on learning (Ahsan et al., 2021).

Normalizer is a method of rescaling each sample that does not have a value of 0 and that has at least one component, independently of other samples, so that its norm is equal to 1.

The above methods were preferred for scaling the data used in the multinomial Naive Bayes and linear support vector classification models, which are used in the study and cannot produce reliable results without scaling. These methods have been tried in algorithms one by one and the most successful methods have been used.

Machine Learning

In this study, machine learning models to be applied to the vectors obtained by Bag of Words and TD-IDF; logistic regression, multinomial Naive Bayes, stochastic gradient descent, linear support vector classifier, perceptron and ridge classifier models. Tree-based classifiers are not preferred because it takes too much time to implement. Since there are approximately 210 thousand data in the data set, models that produce fast results were preferred.

Performance Metrics

Evaluation of the predictions was made by examining the precision (Prec), recall (Rec), f1score (F1), accuracy (Acc) and AUC (Powers, 2011) score. Abbreviations for the required values for the following formulas are: TP = Number of correctly predicted positive labeled samples, FP = Incorrectly predicted positive labeled samples, TN = Correctly predicted negative labeled samples, and FN = Incorrectly predicted negative labeled samples.

RESULTS

In this study, the data mining techniques described above were applied on the “review” feature in the data set, and the results were compared with 5 different feature scaling techniques, 4 different SMOTE methods and 6 different machine learning methods. While the classes have an uneven distribution, feature scaling techniques and machine

learning algorithms are applied and results are obtained without applying SMOTE techniques. The max_features parameter of TF-IDF and bag of words techniques, which is described in the feature extraction section, has been selected as 100,000 and unlimited. Here, the purpose of limiting the max_features parameter is; is the measurement of the relationship between accuracy and time. Learning by including all the features, especially the linear support vector classifier algorithm, requires a lot of hardware power and time. Since a real-time system integration is planned in the future, it is aimed to minimize this time and hardware need by limiting the max_features parameter to 100,000, while making reasonable sacrifices in accuracy.

Some results were obtained from the multinomial Naive Bayes and linear support vector classifier algorithms, which were most affected by the feature scaling techniques described above, and in line with these results, the most successful feature scaling methods were used in the rest of the study. The results are shown in Table 2, 3.

In the TF-IDF method, we can see that for Multinomial Naive Bayes, the best results are obtained with the StandardScaler scaler. In addition, it is seen that MaxAbsScaler and QuantileTransformer scalers are feature scaling methods that give the best results on average for all algorithms.

Before applying the linear support vector classifier algorithm, the algorithm was executed by applying the QuantileTransformer scaler. It is seen that this feature scaling method gives the best results for the linear support vector classifier model with TF-IDF.

In the bag of words method, we can see that for multinomial Naive Bayes, the best results are obtained with the StandardScaler scaler. In addition, it is seen that MaxAbsScaler scaler is the feature scaling method that gives the best results on average for all algorithms.

Before applying the linear support vector classifier algorithm, the algorithm was executed by applying the MaxAbsScaler scaler. It is seen that this feature scaling method gives the best results for the linear support vector classifier model with Bag of Words.

Table 2. NO SMOTE/TF-IDF/max_feature = 100.000

Scale Methods	Model	Class	Precision	Recall	F1 score	Acc	AUC
MaxAbsScaler	MultinomialNB	Negative	0.71	0.74	0.73	0.83	0.81
		Positive	0.89	0.87	0.88		
	LinearSVC	Negative	0.86	0.87	0.86	0.92	0.90
		Positive	0.94	0.94	0.94		
RobustScaler	MultinomialNB	Negative	0.66	0.59	0.62	0.79	0.73
		Positive	0.83	0.87	0.85		
	LinearSVC	Negative	0.87	0.83	0.85	0.91	0.89
		Positive	0.93	0.95	0.94		
Normalizer	MultinomialNB	Negative	0.79	0.41	0.54	0.79	0.68
		Positive	0.79	0.95	0.87		
	LinearSVC	Negative	0.85	0.78	0.81	0.89	0.86
		Positive	0.91	0.94	0.93		
Quantile transformer	MultinomialNB	Negative	0.70	0.76	0.73	0.83	0.81
		Positive	0.89	0.86	0.88		
	LinearSVC	Negative	0.86	0.87	0.87	0.92	0.90
		Positive	0.94	0.94	0.94		
Standard scaler	MultinomialNB	Negative	0.72	0.77	0.74	0.84	0.82
		Positive	0.90	0.87	0.89		
	LinearSVC	Negative	0.79	0.83	0.81	0.89	0.87
		Positive	0.93	0.91	0.92		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 3. No SMOTE/BoW/max_feature = 100.000

Scale methods	Model	Class	Precision	Recall	F1 score	Acc	AUC
MaxAbsScaler	MultinomialNB	Negative	0.72	0.74	0.73	0.84	0.81
		Positive	0.89	0.88	0.88		
	LinearSVC	Negative	0.87	0.87	0.87	0.92	0.90
		Positive	0.94	0.94	0.94		
RobustScaler	MultinomialNB	Negative	0.69	0.75	0.72	0.83	0.80
		Positive	0.89	0.86	0.87		
	LinearSVC	Negative	0.86	0.87	0.87	0.92	0.91
		Positive	0.94	0.94	0.94		
Normalizer	MultinomialNB	Negative	0.86	0.41	0.55	0.80	0.69
		Positive	0.80	0.97	0.88		
	LinearSVC	Negative	0.84	0.78	0.81	0.89	0.86
		Positive	0.91	0.94	0.93		
Quantile transformer	MultinomialNB	Negative	0.70	0.76	0.73	0.83	0.81
		Positive	0.89	0.86	0.88		
	LinearSVC	Negative	0.86	0.87	0.87	0.92	0.91
		Positive	0.94	0.94	0.94		
StandardScaler	MultinomialNB	Negative	0.71	0.76	0.73	0.84	0.81
		Positive	0.90	0.87	0.88		
	LinearSVC	Negative	0.80	0.83	0.81	0.89	0.87
		Positive	0.93	0.91	0.92		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

In line with the results obtained from both tables, StandardScaler scaler will be used for the Multinomial Naive Bayes algorithm in both methods (TF-IDF and BoW) in the machine learning applications to be made from now on. QuantileTransformer scaler will be used for linear support vector classifier algorithm while working in TF-IDF method, MaxAbsScaler scaler will be used for linear support vector

classifier algorithm while operating in Bag of Words method. In this way, the algorithms will work with maximum efficiency.

The results obtained when the data set has uneven class distribution without applying SMOTE techniques are shown in [Tables 4-7](#).

Table 4. No SMOTE/TF-IDF/max_feature no limit

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.85	0.69	0.76	0.87	0.82
	Positive	0.88	0.95	0.91		
Perceptron (MaxAbsScaler)	Negative	0.90	0.86	0.88	0.93	0.91
	Positive	0.94	0.96	0.95		
RidgeClassifier	Negative	0.92	0.84	0.88	0.93	0.90
	Positive	0.93	0.97	0.95		
MultinomialNB (StandarScaler)	Negative	0.84	0.77	0.80	0.89	0.85
	Positive	0.91	0.94	0.92		
SGDClassifier	Negative	0.85	0.66	0.74	0.86	0.92
	Positive	0.87	0.95	0.91		
LinearSVC (StandartScaler)	Negative	0.77	0.77	0.77	0.87	0.84
	Positive	0.90	0.91	0.90		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 5. No SMOTE/BoW/max_feature no limit

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.90	0.87	0.88	0.93	0.91
	Positive	0.94	0.96	0.95		
Perceptron	Negative	0.89	0.88	0.88	0.93	0.92
	Positive	0.95	0.95	0.95		
RidgeClassifier	Negative	0.86	0.86	0.86	0.92	0.90
	Positive	0.94	0.94	0.94		
MultinomialNB (StandarScaler)	Negative	0.85	0.77	0.80	0.89	0.85
	Positive	0.90	0.94	0.92		
SGDClassifier	Negative	0.90	0.88	0.89	0.93	0.97
	Positive	0.95	0.96	0.95		
LinearSVC (StandartScaler)	Negative	0.74	0.78	0.76	0.85	0.83
	Positive	0.91	0.88	0.90		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 6. No SMOTE/TF-IDF/max_feature = 100.000

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.82	0.70	0.76	0.87	0.82
	Positive	0.88	0.94	0.91		
Perceptron (MaxAbsScaler)	Negative	0.86	0.85	0.85	0.91	0.90
	Positive	0.94	0.94	0.94		
RidgeClassifier	Negative	0.86	0.78	0.82	0.90	0.86
	Positive	0.91	0.95	0.93		
MultinomialNB (StandarScaler)	Negative	0.72	0.77	0.74	0.84	0.82
	Positive	0.90	0.87	0.89		
SGDClassifier	Negative	0.83	0.68	0.75	0.86	0.92
	Positive	0.87	0.94	0.91		
LinearSVC (StandartScaler)	Negative	0.79	0.83	0.81	0.89	0.87
	Positive	0.93	0.91	0.92		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 7. No SMOTE/ BoW/max_feature = 100.000

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.87	0.82	0.84	0.91	0.89
	Positive	0.93	0.95	0.94		
Perceptron	Negative	0.85	0.87	0.86	0.92	0.90
	Positive	0.94	0.94	0.94		
RidgeClassifier	Negative	0.80	0.82	0.81	0.88	0.87
	Positive	0.92	0.91	0.92		
MultinomialNB (StandarScaler)	Negative	0.71	0.76	0.73	0.84	0.81
	Positive	0.90	0.87	0.88		
SGDClassifier	Negative	0.88	0.81	0.84	0.91	0.95
	Positive	0.92	0.95	0.94		
LinearSVC (StandartScaler)	Negative	0.79	0.83	0.81	0.89	0.87
	Positive	0.93	0.91	0.92		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 8, 9 were created to observe the SMOTE techniques described in the study and the effect of each technique on the algorithms used. The max_features value has been chosen as 100,000 because of the time it takes to create each SMOTE technique and the high hardware requirement and long waiting times that may occur due to the number of features. Comparing between SMOTE techniques, the most successful SMOTE technique was chosen and results were obtained.

As can be seen from the tables, it is seen that the results of almost all techniques and algorithms are very close. For this reason, it would be the most logical choice to choose the Standard SMOTE technique in terms of time and cost. The SMOTE mentioned in the next results will be the Standard SMOTE technique.

In this section, the following Tables 10-13 were created by using the SMOTE technique, that is, by making the class distribution in the data set more balanced compared to the past.

It is seen in the tables that the linear support vector classifier model gives the best results in almost all results for both methods and the max_features parameter. For the TF-IDF method and the max_features parameter value of 100,000, the best algorithm was the linear support vector classifier algorithm with an Accuracy value of 92%. For the bag of words method and the max_features parameter value of 100,000, the best algorithm was the linear support vector classifier algorithm with an accuracy value of 92%. While the TF-IDF method and max_features parameter had no limit, the best algorithm was the linear support vector classifier

algorithm with an Accuracy value of 93%. While bag of words method and max_features parameter had no limit, the best algorithm was the SGD Classifier algorithm with an Accuracy value of 93%.

DISCUSSION

Although the results obtained from both methods are good, it can be developed by using different methods to create a real-time recommendation system or the existing methods can be improved. Other smote methods such as K-Means smote and SVM smote, which were not used in this study, can be used to minimize the class imbalance problem in the data set. The "usefulCount" property in the dataset can be included in machine learning as an independent variable. In order for bag of words and TF-IDF methods to produce more meaningful word groups, the most appropriate value for the max_features parameter can be determined. In addition to the feature extraction methods used in the study, Word2vec and different feature extraction methods can be used. Also feature extraction can be performed manually. In addition to the machine learning algorithms used in the study, community-based machine learning algorithms can be used. The results can be made even better by making improvements on existing machine learning algorithms.

In this study, sentiment analysis was performed by classification method over categorical text data. The aim is to show the methodology that makes the best use of each data and feature in the data set when performing the sentiment analysis.

Table 8. SMOTE/TF-IDF/max_feature=100.000

SMOTE	Model	Class	Precision	Recall	F1 score	Acc	AUC
Standart smote	LogisticRegression	Negative Positive	0.77 0.91	0.79 0.90	0.78 0.91	0.87	0.85
	Perceptron (MaxAbsScaler)	Negative Positive	0.85 0.94	0.87 0.93	0.86 0.94	0.91	0.90
	RidgeClassifier	Negative Positive	0.82 0.93	0.84 0.92	0.83 0.93	0.90	0.88
	MultinomialNB (StandartScaler)	Negative Positive	0.72 0.90	0.76 0.88	0.74 0.89	0.84	0.82
	SGDClassifier	Negative Positive	0.77 0.90	0.77 0.90	0.77 0.90	0.86	0.92
	LinearSVC (Quantile transformer)	Negative Positive	0.86 0.94	0.87 0.94	0.87 0.94	0.92	0.90
Adasyn	LogisticRegression	Negative Positive	0.77 0.91	0.80 0.90	0.78 0.91	0.87	0.85
	Perceptron (MaxAbsScaler)	Negative Positive	0.85 0.94	0.87 0.94	0.86 0.94	0.92	0.90
	RidgeClassifier	Negative Positive	0.82 0.93	0.85 0.92	0.83 0.93	0.90	0.88
	MultinomialNB (StandartScaler)	Negative Positive	0.72 0.90	0.78 0.87	0.75 0.89	0.84	0.83
	SGDClassifier	Negative Positive	0.76 0.91	0.78 0.90	0.77 0.90	0.86	0.92
	LinearSVC (Quantile transformer)	Negative Positive	0.86 0.94	0.87 0.94	0.87 0.94	0.92	0.91
SMOTE-tomek	LogisticRegression	Negative Positive	0.78 0.91	0.79 0.90	0.78 0.91	0.87	0.85
	Perceptron (MaxAbsScaler)	Negative Positive	0.85 0.94	0.86 0.94	0.85 0.94	0.91	0.90
	RidgeClassifier	Negative Positive	0.82 0.93	0.84 0.92	0.83 0.93	0.90	0.88
	MultinomialNB (StandartScaler)	Negative Positive	0.72 0.90	0.76 0.88	0.74 0.89	0.84	0.82
	SGDClassifier	Negative Positive	0.76 0.90	0.77 0.90	0.77 0.90	0.86	0.92
	LinearSVC (Quantile transformer)	Negative Positive	0.86 0.94	0.87 0.94	0.86 0.94	0.92	0.90
Borderline	LogisticRegression	Negative Positive	0.77 0.91	0.80 0.90	0.78 0.91	0.87	0.84
	Perceptron (MaxAbsScaler)	Negative Positive	0.85 0.94	0.86 0.94	0.86 0.94	0.92	0.90
	RidgeClassifier	Negative Positive	0.82 0.93	0.84 0.92	0.83 0.93	0.90	0.88
	MultinomialNB (StandartScaler)	Negative Positive	0.72 0.90	0.77 0.87	0.74 0.89	0.84	0.82
	SGDClassifier	Negative Positive	0.76 0.91	0.78 0.90	0.77 0.90	0.86	0.92
	LinearSVC (Quantile transformer)	Negative Positive	0.86 0.94	0.87 0.94	0.86 0.94	0.92	0.90

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

CONCLUSION

User comments on various systems we use in our daily life have an important role in choosing the service we want to receive. In this study, in which sentiment analysis was performed on user comments, logistic regression, perceptron, ridge classifier, multinomial Naive Bayes, SGD classifier and linear support vector classifier machine learning algorithms were applied to the data obtained by bag of words and TF-IDF methods. Obtained results were evaluated with precision,

recall, F1 score, accuracy and AUC score metrics. The best result for the bag of words method is the SGD classifier algorithm with an accuracy value of 93%. The best result for the TF-IDF method was the linear support vector classifier algorithm with an accuracy value of 93%. In both methods, there are machine learning algorithms that are very close to the best algorithms and have the same accuracy value. However, when looking at all the evaluation metrics, the best algorithms were selected.

Table 9. SMOTE/BoW/max_feature=100.000

SMOTE	Model	Class	Precision	Recall	F1 score	Acc	AUC
Standart smote	LogisticRegression	Negative	0.83	0.85	0.84	0.90	0.89
		Positive	0.93	0.92	0.93		
	Perceptron	Negative	0.84	0.87	0.86	0.91	0.90
		Positive	0.95	0.93	0.94		
	RidgeClassifier	Negative	0.78	0.84	0.81	0.88	0.87
		Positive	0.93	0.90	0.91		
Adasyn	MultinomialNB (StandartScaler)	Negative	0.71	0.75	0.73	0.84	0.81
		Positive	0.89	0.87	0.88		
	SGDClassifier	Negative	0.82	0.87	0.84	0.90	0.95
		Positive	0.94	0.92	0.93		
	LinearSVC (MaxAbsScaler)	Negative	0.86	0.87	0.86	0.92	0.91
		Positive	0.95	0.94	0.94		
SMOTE-tomek	LogisticRegression	Negative	0.83	0.85	0.84	0.90	0.89
		Positive	0.93	0.93	0.93		
	Perceptron	Negative	0.72	0.82	0.77	0.85	0.84
		Positive	0.92	0.87	0.89		
	RidgeClassifier	Negative	0.78	0.84	0.81	0.88	0.87
		Positive	0.93	0.90	0.91		
Borderline	MultinomialNB (StandartScaler)	Negative	0.71	0.75	0.73	0.84	0.81
		Positive	0.89	0.87	0.88		
	SGDClassifier	Negative	0.84	0.84	0.84	0.90	0.95
		Positive	0.93	0.93	0.93		
	LinearSVC (MaxAbsScaler)	Negative	0.85	0.87	0.86	0.92	0.91
		Positive	0.95	0.94	0.94		
	LogisticRegression	Negative	0.83	0.84	0.84	0.90	0.88
		Positive	0.93	0.93	0.93		
	Perceptron	Negative	0.73	0.82	0.77	0.86	0.85
		Positive	0.92	0.87	0.90		
	RidgeClassifier	Negative	0.78	0.84	0.81	0.88	0.87
		Positive	0.93	0.90	0.91		
	MultinomialNB (StandartScaler)	Negative	0.71	0.75	0.73	0.84	0.81
		Positive	0.89	0.87	0.88		
	SGDClassifier	Negative	0.81	0.88	0.84	0.90	0.95
		Positive	0.95	0.94	0.93		
	LinearSVC (MaxAbsScaler)	Negative	0.85	0.87	0.86	0.92	0.91
		Positive	0.95	0.94	0.94		
	LogisticRegression	Negative	0.83	0.85	0.84	0.90	0.89
		Positive	0.94	0.93	0.93		
	Perceptron	Negative	0.72	0.83	0.77	0.85	0.85
		Positive	0.92	0.86	0.89		
	RidgeClassifier	Negative	0.78	0.84	0.81	0.88	0.87
		Positive	0.93	0.90	0.91		
	MultinomialNB (StandartScaler)	Negative	0.71	0.75	0.73	0.84	0.81
		Positive	0.89	0.87	0.88		
	SGDClassifier	Negative	0.83	0.85	0.84	0.90	0.95
		Positive	0.94	0.93	0.93		
	LinearSVC (MaxAbsScaler)	Negative	0.85	0.87	0.86	0.92	0.91
		Positive	0.95	0.94	0.94		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 10. SMOTE/TF-IDF/max_feature no limit

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.80	0.79	0.80	0.88	0.85
	Positive	0.91	0.92	0.91		
Perceptron (MaxAbsScaler)	Negative	0.90	0.86	0.88	0.93	0.91
	Positive	0.94	0.96	0.95		
RidgeClassifier	Negative	0.90	0.87	0.88	0.93	0.91
	Positive	0.95	0.96	0.95		
MultinomialNB (StandarScaler)	Negative	0.86	0.75	0.80	0.89	0.85
	Positive	0.90	0.95	0.92		
SGDClassifier	Negative	0.78	0.76	0.77	0.87	0.92
	Positive	0.90	0.91	0.90		
LinearSVC (Quantile transformer)	Negative	0.91	0.87	0.89	0.93	0.92
	Positive	0.95	0.96	0.95		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 11. SMOTE/BoW/max_feature no limit

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.87	0.89	0.88	0.93	0.92
	Positive	0.95	0.94	0.95		
Perceptron	Negative	0.86	0.89	0.88	0.93	0.92
	Positive	0.95	0.96	0.95		
RidgeClassifier	Negative	0.83	0.88	0.85	0.91	0.90
	Positive	0.95	0.92	0.93		
MultinomialNB (StandarScaler)	Negative	0.85	0.76	0.80	0.89	0.85
	Positive	0.90	0.94	0.92		
SGDClassifier	Negative	0.88	0.88	0.88	0.93	0.96
	Positive	0.95	0.95	0.95		
LinearSVC (MaxAbsScaler)	Negative	0.87	0.88	0.87	0.92	0.91
	Positive	0.95	0.94	0.95		

Acc: Accuracy

Table 12. SMOTE/TF-IDF/max_feature = 100.000

Model	Class	Precision	Recall	F1 score	Acc	I
LogisticRegression	Negative	0.77	0.79	0.78	0.87	0.85
	Positive	0.91	0.90	0.91		
Perceptron (MaxAbsScaler)	Negative	0.85	0.87	0.86	0.91	0.90
	Positive	0.94	0.93	0.94		
RidgeClassifier	Negative	0.82	0.84	0.83	0.90	0.88
	Positive	0.93	0.92	0.93		
MultinomialNB (StandarScaler)	Negative	0.72	0.76	0.74	0.84	0.82
	Positive	0.90	0.88	0.89		
SGDClassifier	Negative	0.77	0.77	0.77	0.86	0.92
	Positive	0.90	0.90	0.90		
LinearSVC (Quantile transformer)	Negative	0.86	0.87	0.87	0.92	0.90
	Positive	0.94	0.94	0.94		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

Table 13. SMOTE/BoW /max_feature = 100.000

Model	Class	Precision	Recall	F1 score	Acc	AUC
LogisticRegression	Negative	0.83	0.85	0.84	0.90	0.89
	Positive	0.93	0.92	0.93		
Perceptron	Negative	0.84	0.87	0.86	0.91	0.90
	Positive	0.95	0.93	0.94		
RidgeClassifier	Negative	0.78	0.84	0.81	0.88	0.87
	Positive	0.93	0.90	0.91		
MultinomialNB (StandartScaler)	Negative	0.71	0.75	0.73	0.84	0.81
	Positive	0.89	0.87	0.88		
SGDClassifier	Negative	0.82	0.87	0.84	0.90	0.95
	Positive	0.94	0.92	0.93		
LinearSVC (MaxAbsScaler)	Negative	0.86	0.87	0.86	0.92	0.91
	Positive	0.95	0.94	0.94		

SMOTE: Synthetic minority oversampling technique, Acc: Accuracy

ETHICAL DECLARATIONS

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Abbasi Moud, Z., Vahdat-Nejad, H., & Sadri, J. (2021). Tourism recommendation system based on semantic clustering and sentiment analysis. *Expert Syst App*, 167,114324.
- Ahsan, M.M., Mahmud, M.P., Saha, P.K., Gupta, K.D., & Siddique, Z. (2021). Effect of data scaling methods on machine learning algorithms and model performance. *Technologies*, 9(3),52.
- Alamoodi, A.H., Zaidan, B.B., Al Masawa, M., Taresh, S.M., Noman, S., Ahmaro, I.Y., ... & Salahaldin, A. (2021). Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy. *Comput Biol Med*, 139,104957.
- Bao, Y., & Jiang, X. (2016). An intelligent medicine recommender system framework. In 2016 IEEE 11th conference on industrial electronics and applications (ICIEA) (pp. 1383-1388). IEEE.
- Chen, Y., Chang, R., & Guo, J. (2021). Effects of data augmentation method borderline-SMOTE on emotion recognition of EEG signals based on convolutional neural network. *IEEE Access*, 9,47491-47502.
- Dağıstanlı, Ö., Erbay, H., Kör, H., & Yurttakal, A. H. (2023). Reflection of people's professions on social media platforms. *Neural Comput Applications*, 35(7),5575-5586.
- Dinakaran, D., Manjunatha, N., Kumar, C.N., & Math, S.B. (2021). Telemedicine practice guidelines of India, 2020: Implications and challenges. *Indian J Psychiatry*, 63(1),97-101.
- Doulaverakis, C., Nikolaidis, G., Kleontas, A., & Kompatsiaris, I. (2012). GalenOWL: ontology-based drug recommendations discovery. *J Biomed Semantics*, 3,1-9.
- Feng, S., Keung, J., Yu, X., Xiao, Y., & Zhang, M. (2021). Investigation on the stability of SMOTE-based oversampling techniques in software defect prediction. *Informat Software Technol*, 139,106662.
- Fox, S. (2013). Health online 2013. *Pew Intern Am Life Project*.
- Garg, S. (2021). Drug recommendation system based on sentiment analysis of drug reviews using machine learning. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 175-181). IEEE.
- He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In 2008 IEEE International Joint Conference on Neural Networks (pp. 1322-1328). IEEE.
- Hossain, M. D., Azam, M. S., Ali, M. J., & Sabit, H. (2020). Drugs rating generation and recommendation from sentiment analysis of drug reviews using machine learning. In 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE) (pp. 1-6). IEEE.
- Islek, I., & Oguducu, S.G. (2022). A hierarchical recommendation system for e-commerce using online user reviews. *Electronic Commer Res Appl*, 52,101131.
- Jalili, M., Ahmadian, S., Izadi, M., Moradi, P., & Salehi, M. (2018). Evaluating collaborative filtering recommender algorithms: a survey. *IEEE Access*, 6,74003-74024.
- Kaur, H., & Mangat, V. (2017). A survey of sentiment analysis techniques. In 2017 International Conference on I-SMAC. (pp. 921-925). IEEE.
- Kaur, H., & Mangat, V. (2017). A survey of sentiment analysis techniques. In 2017 International Conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC) (pp. 921-925). IEEE.
- Khatte, H., Goel, N., Gupta, N., & Gulati, M. (2021). Movie recommendation system using cosine similarity with sentiment analysis. In 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 597-603). IEEE.
- Oyamada, M. (2019). Extracting feature engineering knowledge from data science notebooks. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 6172-6173). IEEE.
- Pereira, J.G., Tiwari, S., & Ajoy, S. (2020). A survey on filtering techniques for recommendation system. In 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC) (pp. 1-6). IEEE.
- Powers, D.M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv Preprint ArXiv:2010.16061*.
- Pu, W., Liu, N., Yan, S., Yan, J., Xie, K., & Chen, Z. (2007). Local word bag model for text categorization. In Seventh IEEE international conference on data mining (ICDM 2007) (pp. 625-630). IEEE.
- Qaiser, S., & Ali, R. (2018). Text mining: use of TF-IDF to examine the relevance of words to documents. *Int J Comput Appl*, 181(1),25-29.
- Rodríguez, P., Bautista, M.A., Gonzalez, J., & Escalera, S. (2018). Beyond one-hot encoding: lower dimensional target embedding. *Image Vision Computing*, 75,21-31.
- Sharma, A.K., Bajpai, B., Adhvaryu, R., Pankajkumar, S.D., Gordhanbhai, P.P., & Kumar, A. (2023). An efficient approach of product recommendation system using NLP technique. *Mater Today Proceed*, 80,3730-3743.
- Shimada, K., Takada, H., Mitsuyama, S., Ban, H., Matsuo, H., Otake, H., ... & Kaku, M. (2005). Drug-recommendation system for patients with infectious diseases. In AMIA Annual Symposium Proceedings (Vol. 2005, p. 1112). American Medical Informatics Association.
- Tekade, T.N., & Emmanuel, M. (2016). Probabilistic aspect mining approach for interpretation and evaluation of drug reviews. In 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPEs) (pp. 1471-1476). IEEE.

28. Wang, Z.H.E., Wu, C., Zheng, K., Niu, X., & Wang, X. (2019). SMOTETomek-based resampling for personality recognition. *IEEE Access*, 7,129678-129689.
29. Wittich, C. M., Burkle, C. M., & Lanier, W. L. (2014). Medication errors: an overview for clinicians. In Mayo Clinic Proceedings (Vol. 89, No. 8, pp. 1116-1125). Elsevier.
30. Zhang, Y., & Zhang, L. (2022). Movie recommendation algorithm based on sentiment analysis and LDA. *Procedia Comput Sci*, 199,871-878.
31. Zhang, Y., Zhang, D., Hassan, M.M., Alamri, A., & Peng, L. (2015). CADRE: Cloud-assisted drug recommendation service for online pharmacies. *Mobile Networks Appl*, 20,348-355.

Steering angle prediction in autonomous vehicles: a deep learning approach combining VGG16 and LSTM

 Ahmet Mehmet Karadeniz¹,  Nuri Furkan Koçak²

¹Doctoral School of Multidisciplinary Engineering Sciences, Széchenyi István University, Győr, Hungary

²Department of Electronics and Automation, Faculty of Engineering, Kırıkkale University, Kırıkkale, Türkiye

Cite this article: Karadeniz, A.M., & Koçak, N.F. (2024) Steering angle prediction in autonomous vehicles: a deep learning approach combining VGG16 and LSTM. *J Comp Electr Electron Eng Sci*, 2(2), 46-51.

Corresponding Author: Ahmet Mehmet Karadeniz, karadeniz.ahmet.mehmet@sze.hu

Received: 13/09/2024

Accepted: 27/09/2024

Published: 29/10/2024

ABSTRACT

The field of autonomous driving has seen remarkable progress in recent years, with the prediction of steering angles based on varying road conditions emerging as a critical area of focus. While previous efforts have concentrated on lane detection, road object identification, and 3-D reconstruction, our research centers on a vision-based model that leverages deep networks to translate raw camera images into steering angles without the need for predefined feature learning. In our paper, we introduce an end-to-end model that employs deep transfer learning to predict steering angles from image sequences captured by onboard cameras. This model merges two deep learning architectures: a convolutional neural network (CNN) and a long short-term memory (LSTM) network. We utilize the VGG16 network, pre-trained on ImageNet and renowned for its performance, to extract spatial features from the images. Concurrently, the LSTM network processes the temporal information embedded within the image sequences. Our proposed model comprehensively processes spatial-temporal data and adeptly models the nonlinear relationship between the input images and the steering angles. We conducted an experimental study using a publicly available dataset to evaluate the model's effectiveness. The outcomes of our experimental analysis reveal that our model delivers highly efficient and accurate steering angle predictions, effectively emulating human driving patterns. Moreover, our proposed model delivered highly efficient and accurate steering angle predictions, achieving a mean squared error (MSE) of 0.0728 on the validation dataset. This result outperforms conventional models, such as NVIDIA and 3D LSTM, in terms of both accuracy and training efficiency.

Keywords: Steering angle prediction, VGG16, LSTM, deep learning, convolutional neural networks

INTRODUCTION

The advent of autonomous vehicles represents one of the most transformative technological developments of the past decade, with far-reaching implications for road safety, efficiency, and transportation costs. Autonomous vehicles have the potential to significantly reduce traffic accidents, enhance fuel efficiency, and decrease the overall cost of vehicle ownership (Do et al., 2018; Badue et al., 2021). Recent advances in autonomous vehicle technology have been driven by rapid developments in computing power, image processing techniques, and sensor technologies (Velaskar et al., 2014).

Lane departure incidents are predominantly caused by driver distraction and constitute a substantial proportion of vehicular accidents. According to statistics, lane departures were responsible for 51% of traffic accidents in the United States in 2011 (Mammeri et al., 2015). To mitigate these risks, modern vehicles are increasingly being equipped with lane departure warning systems that rely on accurate steering

angle estimation to ensure lane-keeping (Khodayari et al., 2010). The capacity to autonomously perform complex tasks-ranging from driving to surveillance and firefighting-underscores the significance of accurate steering angle prediction (Alshbatat, 2013; Saleem et al., 2021). Accurate steering angle estimation is a critical component of successful lane-keeping and overall vehicular control in these systems. It is recognized as a vital aspect of contemporary vehicle technologies, particularly in the context of developing autonomous driving systems and advanced driver assistance systems (ADAS). Accurate steering angle estimation provides real-time insights into a vehicle's steering dynamics, which is crucial for enhancing road safety and reducing the likelihood of accidents (Aparna et al., 2021).

Approaches to steering angle estimation generally fall into two categories: computer vision-based methods and learning-based methods (Oussama et al., 2020). Computer vision

techniques involve the extraction of features from images to compute the steering angle, whereas learning-based methods utilize neural networks to predict the angle through end-to-end training (Gidado et al., 2020). End-to-end methods have the advantage of direct optimization without having to learn rules and to learn (Jiang et al., 2020). The integration of sophisticated machine learning algorithms, particularly convolutional neural networks (CNNs), has markedly enhanced the capabilities of autonomous driving systems (Song et al., 2022). CNNs have revolutionized pattern recognition by automating the feature extraction process from training data, thereby obviating the need for manually designed feature extraction stages. This automation has led to significant improvements in the accuracy of image-based applications, which are essential for tasks such as lane detection and steering angle prediction (Badue et al., 2021). However, most existing end-to-end steering angle prediction models use a single deep convolutional neural network from sensing to control. These methods can perform well in a single environment, but these models lack memory (Jiang et al., 2020).

In this paper, we propose an innovative steering angle prediction model that combines the Visual Geometry Group 16 (VGG16) model with the long short-term memory (LSTM) to address the drawbacks of existing models. Our model effectively utilizes spatial and temporal information. VGG16 is used to extract spatial features of input images, while LSTM is used to capture temporal dynamics. To reduce the training time, we utilize a transfer learning approach and use the VGG16 model pre-trained on the ImageNet dataset. This method provides faster and more efficient results than training the network from scratch.

Literature Review

Steering angle prediction is a key component in developing autonomous driving systems. VGG16 transfer learning has emerged as a promising approach to improve prediction accuracy while reducing computational complexity. VGG16, a convolutional neural network (CNN) architecture, is particularly effective in image classification tasks, making it suitable for processing visual data from cameras in autonomous vehicles (Song et al., 2022). Using transfer learning, VGG16 can use pre-trained models on large datasets, significantly improving training efficiency and accuracy in steering angle prediction tasks (Alsherif et al., 2023).

To further improve the robustness of the model, data augmentation techniques such as horizontal flipping and random angle adjustments are employed. These strategies diversify the training data set, helping to mitigate overfitting and improve performance in different driving scenarios, as demonstrated in road tests (Song et al., 2022). In addition, the integration of VGG16 with LSTM networks allows the model to capture temporal dependencies in sequential image data, leading to more accurate steering predictions. By integrating the dynamics of vehicle control over time, this approach improves steering angle prediction accuracy by approximately 95%, outperforming traditional CNN models (Hoang et al., 2023).

Despite the advantages of VGG16 and transfer learning, there are still challenges in real-time processing and adaptation to different driving environments. These limitations highlight the need for further research and development in this area

to improve the practical application of these models in autonomous vehicles.

In particular, the transfer learning approach with VGG16 is particularly beneficial in scenarios where labeled data is limited. By fine-tuning VGG16 to specific driving datasets, researchers can achieve improved steering angle prediction as the model retains learned features from large image datasets. The depth and architecture of VGG16 allow it to capture intricate patterns in visual data, which is critical for understanding and predicting driving scenarios. In addition, transfer learning reduces the dependence on large datasets, making it possible to train models in environments where data collection is difficult (Ismail et al., 2024).

Studies have shown that models using VGG16 for steering angle prediction can achieve high accuracy, with performance metrics such as precision and recall significantly improved over traditional methods (Golnari et al., 2024; Karadeniz et al., 2024). However, it is important to explore alternative architectures such as DenseNet or ResNet, which may offer competitive performance and efficiency in certain contexts (Ismail et al., 2024).

In this study we have gone through the following questions and answered them in Section 3:

RQ1: How does combining VGG16 with LSTM improve steering angle prediction?

RQ2: What are the key factors that contribute to the accuracy and efficiency of the model?

RQ3: What are the potential limitations of the model in different driving environments?

METHODS

Convolutional Neural Networks

Over recent years, it has become clear that convolutional neural networks (CNNs) are highly effective at generating a detailed representation of an input image by transforming it into a fixed-length vector. This representation is versatile and can be applied to a range of visual tasks. CNNs excel at processing visual imagery. Typically, a CNN is composed of an input layer, an output layer, and several hidden layers. The hidden layers include a sequence of convolutional layers, pooling layers, normalization layers, and fully connected layers. The convolutional layers perform operations on the data (often images) using multiplication or a dot product. Although these operations are traditionally called convolutions, they are mathematically carried out by a sliding window or filter executing a dot product with the image. This process is crucial for determining how weights are assigned at specific index points within the matrix (Turk, 2024). The resulting output from the convolutional layer is passed on to the subsequent pooling layer, whose role is to downsize the data to simplify the computation. Owing to their exceptional ability to interpret visual data, CNNs and their variations are now being utilized for predicting steering angles. Researchers have investigated various CNN architectures, which differ in the number and size of layers and neurons, to optimize steering angle prediction (Song et al., 2022; Turk, 2024).

LSTMs, or long short-term memory networks, differ from traditional feedforward neural networks by having feedback connections. This allows LSTMs to not only process

individual data points, like images, but also to handle entire sequences of data, such as speech or video streams. LSTMs have demonstrated superior capabilities in learning and understanding sequences, which makes them suitable for autonomous vehicles (AVs) to maintain lane positions by analyzing the relationship between consecutive image frames. Typically, LSTMs are utilized after a CNN, which is responsible for feature detection, to comprehend temporal relationships within the data. This combined CNN-LSTM model has been adopted in various research studies (Saleem et al., 2021).

Data Processing

The dataset used in this study can be found in Kaggle as shown in. It includes approximately 10000 images along with their associated steering angle data. Figure 1 showcases a histogram of the collected data, depicting the distribution of steering angles against the frequency of their occurrence in each bin.

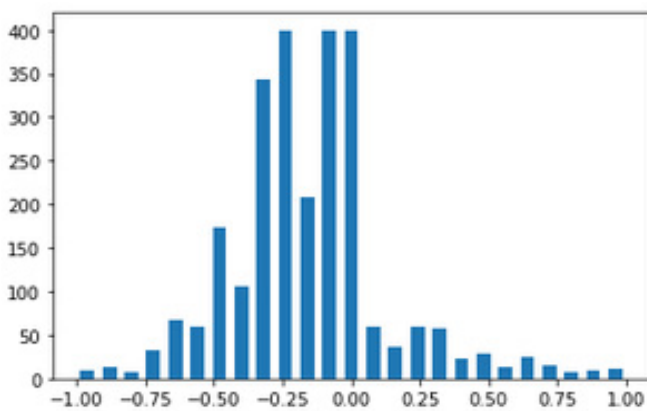


Figure 1. Steering angles and their frequencies

After collecting the driving data, it undergoes several preprocessing steps to make it suitable for training purposes. These steps involve a range of techniques such as data cleansing, image cropping and resizing as well as removing superfluous data and dealing with any missing values. The camera images are cropped to focus on the road lane markings which are the area of interest. Then, these cropped images are resized to conform to the specific dimensions required for the model's input. Preprocessing the images facilitates faster processing and ensures the model focuses on extracting only the relevant datasets. Figure 2, 3 illustrate the images from the front-facing camera before the preprocessing steps and after, respectively.

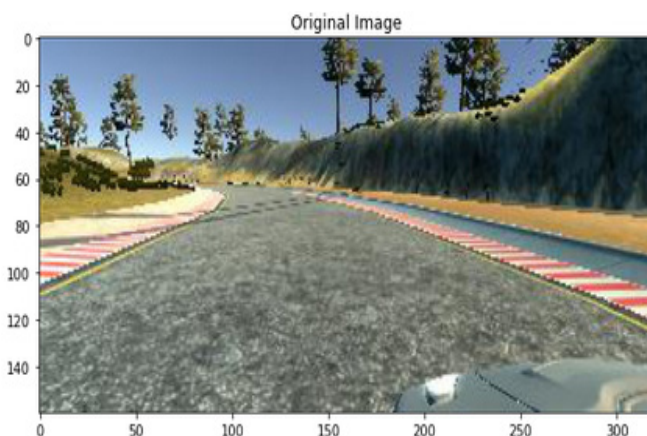


Figure 2. Original image without preprocessing

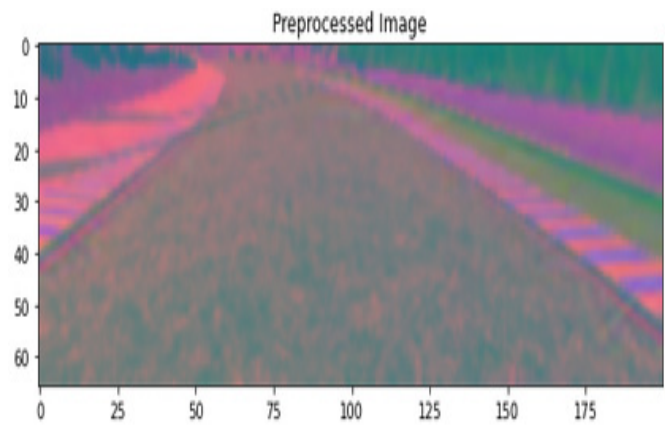


Figure 3. Preprocessed image

After preprocessing, the data is divided into two distinct subsets: 80% of it for the training and 20% for the validation. The training dataset is used to train the deep learning model and the validation dataset allows for performance evaluation during the training process and also it is used to evaluate the model's final performance. This split is a standard deep learning technique for preventing overfitting and underfitting. The Sci-kitlearn library has been utilized to randomly partition the data using Pandas and NumPy tools. Figure 4 illustrates the common steering angles across the training and validation datasets.

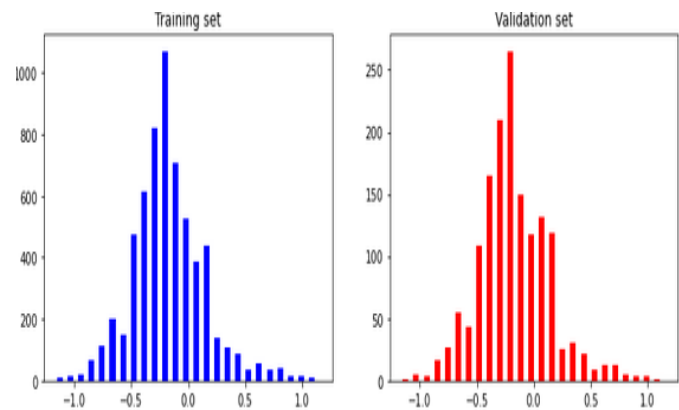


Figure 4. Training and validation datasets

Model Architecture

The next phase involves defining the model architecture. For this purpose, the VGG16 transfer learning model, which has been pre-trained on the ImageNet dataset and is available in the Tensorflow. Keras library, is chosen. The base layers of the VGG16 are frozen to inhibit any additional training. To customize the model for the steering angle prediction task, custom layers are integrated on top of the VGG16 base. An LSTM layer is added to process sequential data and capture temporal dependencies, which is crucial for understanding sequences like video frames. This is followed by several dense layers, including fully connected layers with 100 neurons, 50 neurons, 10 neurons, and finally 1 neuron to output the predicted steering angle. These layers employ as the activation function an exponential linear unit (ELU) to introduce nonlinearity. The culmination of the model architecture is a solitary neuron that serves as the output layer, tasked with predicting the steering angle. The structural designs of the model, including the integration of VGG16 and LSTM layers as well as number of parameters, are depicted in Figure 5.

Layer (type)	Output Shape	Param #
time_distributed_8 (TimeDist (None, 5, 2, 6, 512))		14714688
time_distributed_9 (TimeDist (None, 5, 1, 3, 24))		387224
time_distributed_10 (TimeDis (None, 5, 1, 2, 36))		21636
time_distributed_11 (TimeDis (None, 5, 1, 1, 48))		43248
time_distributed_12 (TimeDis (None, 5, 1, 1, 64))		27712
time_distributed_13 (TimeDis (None, 5, 1, 1, 64))		36928
time_distributed_14 (TimeDis (None, 5, 1, 1, 64))		0
time_distributed_15 (TimeDis (None, 5, 64))		0
lstm_2 (LSTM)	(None, 5, 100)	66000
dropout_5 (Dropout)	(None, 5, 100)	0
lstm_3 (LSTM)	(None, 50)	30200
dropout_6 (Dropout)	(None, 50)	0
dense_2 (Dense)	(None, 10)	510
dropout_7 (Dropout)	(None, 10)	0
dense_3 (Dense)	(None, 1)	11
Total params: 15,248,157		
Trainable params: 533,469		
Non-trainable params: 14,714,688		

Figure 5. Detailed architecture of the proposed steering angle prediction model, combining VGG16 for spatial feature extraction and LSTM for temporal dependency handling

VGG16: Visual Geometry Group 16, LSTM: Long short-term memory

Evaluation Metrics and Training Process

During the training phase, the preprocessed data is input into the model for learning. The TensorFlow and Keras frameworks facilitate the training of the models. Throughout this phase, hyperparameters such as batch size and learning rate are fine-tuned. As a loss function mean squared error (MSE) is chosen. Then the Adam optimizer is leveraged to reduce the MSE loss, which quantifies the difference between the predicted steering angles and the actual values. Key performance indicators, specifically training and validation loss are monitored to gauge the model's effectiveness during the training. Upon completion of the training, the model is saved for future use in determining the steering angles for the autonomous vehicle. To capture the best model weights, a checkpoint callback function is utilized which ensures the model is saved at the point of lowest loss throughout the epochs.

The effectiveness of the models is evaluated using the test dataset. The MSE is used as the metric for evaluation. The optimal value for MSE is zero, indicating perfect predictions with no errors. The mathematical expression for calculating MSE is presented in equation (1).

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

Where:

n is the number of samples,

$\hat{y}_i$ represents the predicted steering angle,

y_i represents the actual steering angle,

Σ indicates summation over all samples.

The mean squared error (MSE) is selected as the primary evaluation metric due to its sensitivity to large errors, which is crucial in steering angle prediction. Large deviations in steering angles can lead to lane departure or accidents, making MSE an effective measure of model performance. Furthermore, MSE penalizes larger errors more heavily, ensuring that the model focuses on minimizing significant prediction deviations. The difference between training and validation losses indicates the model's generalization ability. In our experiments, the small gap between these losses demonstrates that the model did not overfit to the training data, validating its robustness across different datasets.

RESULTS AND DISCUSSION

In our study, we utilized a Kaggle notebook environment consisting of GPU P100 accelerator and Keras for training and validating our neural networks. We used a dataset from Kaggle and provided it in, comprising 1802 training samples and 451 validation samples due to limited computational resources. During the training phase, the model underwent 20 epochs with a learning rate set at 0.001 with Adam optimizer and a batch size of 32. The losses of MSE score per epochs are demonstrated in Figure 6. Remarkably, the training was completed in just 41 minutes, demonstrating a more efficient use of time compared to other models.

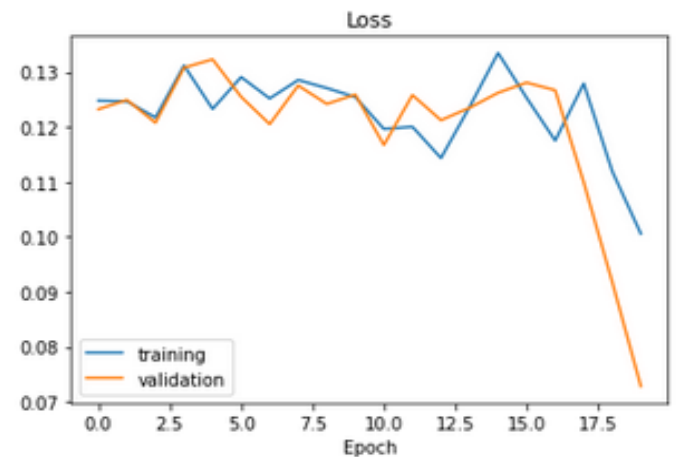


Figure 6. Training and validation losses per epochs

The robustness of our model is further enhanced through data augmentation techniques such as horizontal flipping and random angle adjustments. These methods ensure the model can generalize well across different driving scenarios, reducing the risk of overfitting. Furthermore, the integration of temporal dependencies via LSTM networks allows the model to handle sequential data more effectively, improving performance under real-world conditions.

Limitations

Also, the proposed model combining VGG16 and LSTM resulted 0.0728 MSE score in the validation dataset and its comparison with other relevant studies are given in Table. According to the comparison, our model surpassed the NVIDIA and 3D LSTM, however performed slightly worse against VGG16. The reason of it might be because of the used dataset size, computational resources such as accelerators (GPU) or the limitation of the dataset used. For example, if a dataset is chosen which includes not diverse values, then obviously the loss will be much lower.

Table. Comparison of the models

Model	MSE score
NVIDIA (Du et al., 2019)	0.3154
3D LSTM (Du et al., 2019)	0.3374
VGG16 (Jiang et al., 2020)	0.0413
Our proposed model	0.0728

MSE: Mean squared error, VGG16: Visual Geometry Group 16, LSTM: Long short-term memory

CONCLUSION

In this study, we introduce an End-to-End steering angle prediction model that fuses a modified VGG16 network with LSTM to effectively harness both spatial and temporal information from input image sequences. The modified VGG16, based on transfer learning, allows for efficient training, while the LSTM network captures crucial temporal data. Our mixed network design, validated on the publicly available dataset demonstrates accurate steering angle predictions and robustness across different environments, as evidenced by our model's lower loss value in terms of MSE score for the transfer learning and 3D convolutional models.

The visualization of the convolutional layers reveals that our model can learn key road features without predefined parameters an advantage of the End-to-End approach. Despite the promising results, these models have not yet been tested in real-world conditions or specialized datasets, an endeavor we aim to undertake in future work. We believe the robustness shown in experimental results bodes well for real-time autonomous driving applications.

Looking ahead, we see potential in expanding the deep learning model's architecture with larger and deeper layers, which may yield improved outcomes despite current computational limitations. However, there remains significant research to be done before such models can be widely deployed. Future models could benefit from training on more diverse data. Additionally, a high-quality simulator coupled with deep reinforcement learning, guided by a reward function prioritizing efficiency, ride smoothness, rule adherence, and safety, could further refine these models for public transportation use.

Benefits and Drawbacks

Every approach has advantages and disadvantages. We also went through analysis of the benefits and drawbacks of our approach and they can be found as the following:

Benefits: The combination of VGG16 and LSTM captures both spatial and temporal features, leading to improved steering angle prediction accuracy.

Transfer learning with VGG16 significantly reduces training time.

The model outperforms other architectures like NVIDIA and 3D LSTM in certain scenarios.

Drawbacks: The model has not been tested in real-world driving environments, and the dataset used may not represent all possible driving conditions.

Computational resource limitations restricted the dataset size and could have impacted model performance.

ETHICAL DECLARATIONS

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Do, T.D., Duong, M.T., Dang, Q.V., & Le, M.H. (2018). Real-time self-driving car navigation using deep neural network. In 2018 4th International Conference on Green Technology and Sustainable Development (GTSD) (pp. 7-12). IEEE. doi:10.1109/GTSD.2018.8595590
- Badue, C., Guidolini, R., Carneiro, R.V., Azevedo, P., Cardoso, V.B., Forechi, A., ... & De Souza, A.F. (2021). Self-driving cars: a survey. *Expert Systems Appl*, 165,113816. doi:10.1016/j.eswa.2020.113816
- Velaskar, P., Vargas-Clara, A., Jameel, O., & Redkar, S. (2014). Guided navigation control of an unmanned ground vehicle using global positioning systems and inertial navigation systems. *Int J Electr Comput Engineer*, 4(3),2088-8708. doi:10.11591/ijece.v4i3.5183
- Mammeri, A., Lu, G., & Boukerche, A. (2015). Design of lane keeping assist system for autonomous vehicles. In 2015 7th International Conference on New Technologies, Mobility and Security (NTMS) (pp. 1-5). IEEE. doi:10.1109/NTMS.2015.7266483
- Khodayari, A., Ghaffari, A., Ameli, S., & Flahatgar, J. (2010). A historical review on lateral and longitudinal control of autonomous vehicle motions. In 2010 International Conference on Mechanical and Electrical Technology (pp. 421-429). IEEE. doi:10.1109/ICMET.2010.5598396
- Alshbatat, A.I.N. (2013). Behavior-based approach for the detection of land mines utilizing off the shelf low cost autonomous robot. *Int J Robotics Automation*, 2(3),83. doi:10.11591/ijra.v2i3.2038
- Saleem, H., Riaz, F., Mostarda, L., Niazi, M.A., Rafiq, A., & Saeed, S. (2021). Steering angle prediction techniques for autonomous ground vehicles: a review. *IEEE Access*, 9,78567-78585. doi:10.1109/ACCESS.2021.3083890
- Aparna, M.P., Gandhiraj, R., & Panda, M. (2021). Steering angle prediction for autonomous driving using federated learning: The impact of vehicle-to-everything communication. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-7). IEEE. doi:10.1109/ICCCNT51525.2021.9580097
- Oussama, A., & Mohamed, T. (2020). A literature review of steering angle prediction algorithms for self-driving cars. *Advan Intell Syst Appl Comput Sci*, 4,30-38. doi:10.1007/978-3-030-36674-2_4
- Gidado, U.M., Chiroma, H., Aljojo, N., Abubakar, S., Popoola, S.I., & Al-Garadi, M.A. (2020). A survey on deep learning for steering angle prediction in autonomous vehicles. *IEEE Access*, 8,163797-163817. doi:10.1109/ACCESS.2020.3017883
- Jiang, H., Chang, L., Li, Q., & Chen, D. (2020). Deep transfer learning enable end-to-end steering angles prediction for self-driving car. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 405-412). IEEE. doi:10.1109/IV47402.2020.9304611
- Song, X., Cao, H., & You, H. (2022, December). Steering Wheel Rotation Angle Prediction Based on VGG-16 and Data Augmentation. In 2022 4th International Conference on Control and Robotics (ICCR) (pp. 398-402). IEEE. doi:10.1109/ICCR55715.2022.10053867
- Alsherif, M., Daowd, M., Bassiuny, A.M., & Metered, H.A. (2023). Utilizing transfer learning in the udacity simulator to train a self-driving car for steering angle prediction. In 2023 Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS) (pp. 134-139). IEEE. doi:10.1109/ICICIS58388.2023.10391185

14. Song, X., Cao, H., & You, H. (2022). Steering wheel rotation angle prediction based on VGG-16 and data augmentation. In 2022 4th International Conference on Control and Robotics (ICCR) (pp. 398-402). IEEE. doi:10.1109/ICCR55715.2022.10053867
15. Hoang, T.N., Hong, P.P., Vinh, N.N., Nguyen, N.T., Nguyen, K.H., & Quach, L.D. (2023). An improved lane-keeping controller for autonomous vehicles leveraging an integrated CNN-LSTM approach. *Int J Advan Comput Sci Appl*, 14,7. doi:10.14569/IJACSA.2023.0140723
16. Ismail, S.N.M.S., Razak, S.F.A., & Aziz, N.A.A. (2024). Transfer learning for improved electrocardiogram diagnosis of cardiac disease: exploring the potential of pre-trained models. *Bullet Electr Engineer Informat*, 13(5),3288-3300. doi:10.11591/eei.v13i5.7053
17. Ismail, S.N.M.S., Razak, S.F.A., & Aziz, N.A.A. (2024). Transfer learning for improved electrocardiogram diagnosis of cardiac disease: exploring the potential of pre-trained models. *Bullet Electr Engineer Informat*, 13(5),3288-3300. doi:10.11591/eei.v13i5.7053
18. Golnari, A., Komeili, M.H., & Azizi, Z. (2024). Probabilistic deep learning and transfer learning for robust cryptocurrency price prediction. *Expert Syst Appl*, 124404. doi:10.1016/j.eswa.2024.124404
19. Karadeniz, A. M., Ballagi, Á., & Kóczy, L. T. (2024). Transfer learning-based steering angle prediction and control with fuzzy signatures-enhanced fuzzy systems for autonomous vehicles. *Symmetry*, 16(9),1180. doi:10.3390/sym16091180
20. Turk, F. (2024). Investigation of machine learning algorithms on heart disease through dominant feature detection and feature selection. *SIViP*, 18,3943-3955. doi:10.1007/s11760-024-03060-0
21. Turk F. (2024). RNGU-NET: a novel efficient approach in segmenting tuberculosis using chest X-ray images. *PeerJ Comput Sci*, 10:e1780. doi:10.7717/peerj-cs.1780
22. <https://www.kaggle.com/datasets/zaynena/selfdriving-car-simulator> (accessed on 25 August 2024).
23. Du, S., Guo, H., & Simpson, A. (2019). Self-driving car steering angle prediction based on image recognition. arXiv preprint arXiv:1912.05440. doi:10.48550/arXiv.1912.05440

Determination of I-V curves for photovoltaic generator by an analytical method: Akbaba Model

 Nurettin Gökşenli,  Enes Bektaş,  Mehmet Tümay

Department of Electronics and Automation, Faculty of Engineering, Çankırı Karatekin University, Çankırı, Türkiye

Cite this article: Gökşenli, N., Bektaş, E., & Tümay, M. (2024) Determination of I-V curves for photovoltaic generator by an analytical method: Akbaba Model. *J Comp Electr Electron Eng Sci*, 2(2), 52-55.

Corresponding Author: Nurettin Gökşenli, ngoksenli@karatekin.edu.tr

Received: 29/07/2024

Accepted: 07/10/2024

Published: 29/10/2024

ABSTRACT

In order to maximize the amount of electrical energy generated and ensure effective use, the conventional I-V or V-I characteristics derived from the diode model of photovoltaic (PV) solar panels are important. Unfortunately, these features cannot be extracted using linear equations. The conventional I-V or V-I characteristics requires extremely drawn-out, time-consuming procedures. Many studies focus on calculating maximum power extraction. The analytical solutions to these equations are also nonexistent. Programming on computers can assist in solving the equations. The Akbaba Model was created specifically for the photovoltaic generator (PVG) to address these challenges. The Akbaba Model relies heavily on coefficients A, B, and C in this equation. The diode model serves as the true model in this case, and the coefficients A, B, and C are calculated. Another advantage of this model is that, because its qualities have partially deteriorated with time, the genuine I-V characteristics of the PV panels can be recovered by measuring again.

Keywords: Electrical power generation, maximum power, photovoltaic, renewable energy, solar panel

INTRODUCTION

The amount of energy consumed in the world is increasing day by day due to reasons such as global growth, the rapid increase in technology and industrial development, and therefore the way of life keeping up with this. However, the trend towards new energy sources is increasing day by day due to reasons such as traditional energy sources having limited reserves, being difficult and expensive to find and use, and causing environmental pollution. The alternative to this is renewable energy sources, which are both environmentally friendly and generally have no reserve limit. One of these, perhaps the most important; it is the sun. Because it does not pollute the environment, can be used anywhere, and is an endless source of energy. Photovoltaic (PV) solar panels are used to convert solar energy into electrical energy. PV solar panels are among the renewable energy technologies that have found the greatest development and application area. PV solar panel is a photovoltaic generator (PVG). PVs have very low maintenance costs and important advantages such as not requiring fuel, not creating pollution and not making noise. The fact that silicon, the primary raw material used in PV cells, is widely available in nature is another significant benefit. In addition, we can say that its disadvantages are high investment costs and low energy conversion efficiency. For this reason, studies to improve the efficiency of PV solar panels continue intensively (NREL).

To improve the efficiency of PV solar panels, studies are carried out either to increase the efficiency of PV solar cells or to use electrical energy at the maximum level of the output of PV solar panels (Akbaba, 2006; Brambilla; 1999, Kulaksiz, 2012). This study is about the second method. In other words, it is the study of using the electrical energy at the maximum level at the output of PV solar panels. In the literature, two ways are followed to convert solar energy into electrical energy using solar panels. These are maximum power point tracker (MPPT) (Ali et al., 2020; Motahhir et al., 2020; Yang et al., 2020; Shankar & Saravana, 2020; Karami et al., 2017; Rezk et al., 2017) and solar tracking systems methods (Hasanah et al., 2020; Zegrar et al., 2021; De Riso et al., 2024). There are many studies in the literature on these methods in different fields (Akbaba, 2007; Appelbaum & Sarma, 1989; Saied, 1988; López-Lapeña et al., 2009). It has been observed that the amount of solar radiation increases by 30% in single-axis solar tracking systems and 35% in dual-axis solar tracking systems compared to fixed solar panels. In MPPT systems, this ratio is much higher. Because MPPT systems provide maximum electrical power at the output by converting the maximum amount of solar radiation into electrical energy despite different environmental conditions (Akbaba, 2007). Machine learning algorithms and artificial intelligence are used in the literature (Türk, 2024; Turk, 2024).

Similarly (Voutsinas et al., 2023; González et al., 2021; Sardar et al., 2024; Hocine et al., 2021; Oliva et al., 2014), emphasize the necessity of evaluating V-I characteristics to improve PV system efficiency, fault detection, and overall performance. They use innovative techniques such as machine learning, imaging, and optimization algorithms to improve the dependability of solar energy generation (Voutsinas et al., 2023). Investigates the use of machine learning techniques for defect detection in PV systems using I-V curve data. Support vector machines (SVMs) and deep neural networks are used to detect typical problems such as line-to-line and ground faults. The work focuses on the accuracy of ML models in diagnosing defects in real-time systems, which improves total PV system dependability (González et al., 2021). focuses on creating an IoT-based I-V curve tracer for real-time monitoring of solar panels. The tracer allows for extensive investigation of V-I characteristics, which improves PV array performance and problem identification while also giving vital data for operational modifications (Sardar et al., 2024). focuses on how series resistance impacts the efficiency of silicon solar cells. The researchers use advanced imaging techniques to assess the relationship between resistance, voltage, and current in order to optimize cell design. The I-V curve is critical in determining the energy losses due by excessive resistance in cells. In (Hocine et al., 2021), a combination of thermal imaging and I-V curve analysis is used to detect flaws in PV modules, such as microcracks and hot spots. This technique allows for early detection of performance decline, especially under changing temperature conditions (Oliva et al., 2014). describes the artificial bee colony (ABC) algorithm for optimizing the I-V properties of PV cells. The method accurately predicts the maximum power point (MPP) under changing temperatures and irradiance, increasing the total efficiency of PV systems.

Methodology

The analysis of the effectiveness of systems powered by PV generators depends heavily on the mathematical model of the electrical properties of PV generators. Because it must be used in design for maximum energy use. Therefore, the mathematical model of the I-V characteristic is one of the most important key elements in the performance analysis of loads operated with PV generators (Akbaba, 2007). All published literature utilizes a consistent mathematical model to describe the current-voltage (I-V) or voltage-current (V-I) characteristics of PV cells or PV cell generators. This model is based on the traditional p-n junction, which has been widely employed in electronic devices for a long time. The I-V characteristic model accurately captures the relationship between the current output and the voltage of PV cells under various conditions, providing crucial insights into their performance and efficiency (Appelbaum & Sarma, 1989).

I can be formulized as given in equation (1). Traditional I-V characteristic of solar panels is obtained by using equation (1).

$$I = I_{ph} - I_0 [e^{L(V+IR_s)} - 1] \quad (1)$$

And, to obtain V-I characteristic, equation (1) can be written as in equation (2) by leaving V alone;

$$V = IR_s + \frac{1}{L} \ln \left[\frac{I_{ph} - I + I_0}{I_0} \right] \quad (2)$$

It is shown with equations. Here, I_{ph} is produced light current, I_0 is reverse saturation current, L is PV generator constant, and R_s is series resistance of the PV generator. The equations given above with the traditional I-V and V-I characteristics are quite non-linear. As a result, the solution requires difficult, extended, and complex methods. The following model, referred to as the Akbaba Model in the PVG literature, is proposed in order to eliminate suggested problems (Akbaba & Alattawi, 1995).

The current, can be formulized by Akbaba Model as in equation (3) as following:

$$I = \frac{V_{oc} - V}{A + BV^2 - CV} \quad (3)$$

The Akbaba Model relies heavily on coefficients A, B, and C. The real diode model is utilized to obtain the A, B, and C coefficients.

It is found with $A = \frac{V_{oc}}{I_{sc}}$. Here, V_{oc} is the open circuit voltage and I_{sc} is the short circuit current. The other two coefficients will be calculated using the values at point "a", which is on the I-V Characteristic and corresponds to 94% of the short circuit current, and point "b", which corresponds to 68% (Akbaba & Alattawi, 1995). B and C coefficients can be obtained by using the mathematical operations as follows.

$$B = \frac{K_1 - K_2}{K_3} \quad (4)$$

$$C = \frac{K_1 V_a - K_2 V_b}{K_3} \quad (5)$$

where,

$$K_1 = V_a I_a (V_{oc} - V_b - A I_b) \quad (6)$$

$$K_2 = V_b I_b (V_{oc} - V_a - A I_a) \quad (7)$$

$$K_3 = V_a I_a V_b I_b (V_b - V_a) \quad (8)$$

For the calculation of PVG power, to be used in $P=VI$ and the power is calculated by equation (9).

$$P = V(V_{oc} - V) / (A + BV^2 - CV) \quad (9)$$

By taking the derivative of equation (9), equation (10) is obtained for the maximum point with $\frac{dP}{dV} = 0$.

$$V_{max}^2 - Q_1 V_{max} + Q_2 = 0 \quad (10)$$

And, maximum voltage V_{max} is formulized by solving equation (10);

$$V_{max} = 0,5 \left(Q_1 - \sqrt{Q_1^2 - 4Q_2} \right) \quad (11)$$

where,

$$Q_1 = \frac{2A}{C - BV_{oc}} \quad (12)$$

$$Q_2 = \frac{AV_{oc}}{C - BV_{oc}} \quad (13)$$

To obtain the maximum current equation from here, if we write the V_{max} value instead of V in the Akbaba Model equation in equation (3), the maximum current is obtained as in equation (14).

$$I_{max} = \frac{V_{oc} - V_{max}}{A + BV_{max}^2 - CV_{max}} \quad (14)$$

As a result, maximum power can be calculated with the equation $P_{\max} = V_{\max} I_{\max}$.

Thus, I-V characteristic curves of PV generators can be easily drawn by analytical analysis method using the Akbaba model. For this reason, the method presented as the Akbaba Model is a new and simpler, but accurate model for the I-V characteristic curve (Goksenli & Akbaba, 2016). The accuracy, flexibility, and simplicity of this new model have been proven by comparing it with the traditional model. In (Akbaba & Alattawi, 1995), Akbaba Model is proposed in the PV generators literature, by comparing it with the traditional model in a simulation environment, they proved that there is a negligible difference of around 1-2% between the traditional and the Akbaba model.

Application Method and I-V Plotting the Characteristic Curve

In this study, measurements of PV solar panels at different solar radiations are made for the first time, calculations are made, and the I-V characteristic curve is drawn analytically. Another advantage of this study is that since some electrical properties of the solar panels used may change over time, this method allows the I-V characteristic curve to be drawn by re-measurement.

To draw the I-V characteristic curve of the solar panel, the V_{oc} (open circuit voltage), I_{sc} (short circuit current) of the solar panel, and the current and voltages at different loads were measured under real solar radiation by connecting a rheostat to the output. The values obtained as a result of measurements at different solar irradiance values were substituted into the Akbaba Model equation (3), calculations were made and compared with the measurement results, and very close results were obtained. These results showed insignificant differences of around 1-2 percent, as in the study simulated in the literature (Akbaba & Alattawi, 1995). Some of the data obtained as a result of measurements and calculations are shown in Table. In Table, some of the current (I) and voltage (V) values from numerous measurements at different solar irradiances are given in separate columns.

Table. I and voltage V values measured at different solar radiation

Solar irradiance 35%		Solar irradiance 50%		Solar irradiance 85%		Solar irradiance 100%	
V_1 (Volt)	I_1 (Amper)	V_2 (Volt)	I_2 (Amper)	V_3 (Volt)	I_3 (Amper)	V_4 (Volt)	I_4 (Amper)
0.00	0.35	0	0.50	0.00	0.59	0.00	0.68
3.00	0.35	5.55	0.50	2.00	0.59	1.00	0.68
4.78	0.35	8.15	0.49	3.00	0.59	2.00	0.68
6.30	0.35	11.06	0.49	4.00	0.59	3.00	0.68
8.95	0.35	12.10	0.49	5.00	0.59	4.00	0.68
10.20	0.35	13.10	0.49	6.00	0.59	5.00	0.68
11.00	0.35	14.00	0.48	7.00	0.59	6.00	0.68
13.12	0.34	14.85	0.47	8.00	0.59	7.00	0.68
14.18	0.33	15.37	0.45	8.50	0.59	8.00	0.68
15.27	0.31	15.72	0.44	9.00	0.59	8.79	0.68
15.95	0.28	16.05	0.42	9.70	0.58	10.57	0.68
16.48	0.25	16.32	0.40	10.33	0.58	11.83	0.68

I: Current, V: Voltage

By using the data obtained as a result of numerous measurements made at different solar radiations, the I-V characteristic curves of the PV solar panel were calculated analytically and easily drawn with the Akbaba Model equation (3) (Figure). Thus, there is no need for the tedious, time-consuming, and difficult calculations required in equations (1 and 2), which are quite non-linear and are traditionally used in I-V characteristic drawings for PV solar panels. Figure shows the I-V Characteristic curves plotted at different percent solar irradiances according to the measurements made in Table. Thus, the I-V characteristic curve according to different solar radiation in Table is plotted as in the literature.

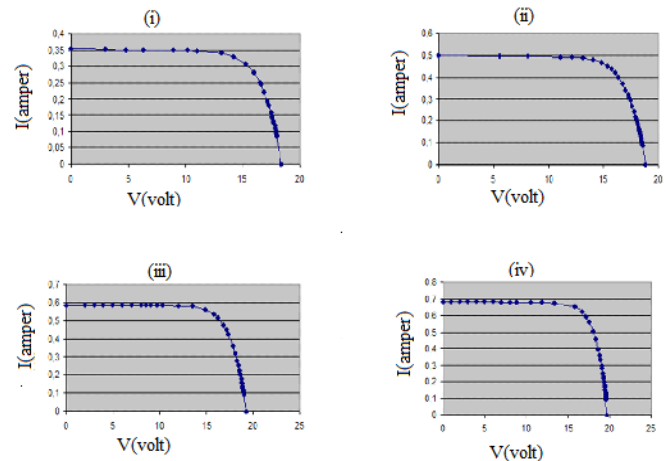


Figure. I-V Characteristic curves plotted at different percent solar irradiances; i) Radiation 35% ii) Radiation 50% iii) Radiation 85% iv) Radiation 100%

Defining the I-V curves of PV panels involves measuring current and voltage at various loads, typically using a solar simulator or real sunlight. Engineers adjust the load to record data points, creating the I-V curve, which shows how the panel performs under different conditions. In real-world applications, engineers use I-V curves to assess the panel's efficiency, identify its MPP, and ensure optimal operation in solar installations. This helps in system design, fault detection, and maximizing energy output in varying environmental conditions.

CONCLUSION

Traditional I-V or V-I characteristics determined from the diode model of PV solar panels are critical in determining maximum power points and utilizing the generated electrical energy efficiently. Because these features reveal the point of maximum energy. However, the equations commonly utilized to derive these features are non-linear. As a result, the solution necessitates lengthy, time-consuming processes. Furthermore, these equations lack analytical solutions. Thus, equations can be solved using computer programs. To address these issues, the Akbaba Model has been used. This approach is simpler and better suited for application and calculation than standard equations. Most importantly, in our study, it has been shown that this model can also be examined analytically.

Another advantage of the suggested model is that due to the partial loss of electrical properties in PV solar panels over time, the true I-V characteristics of PV solar panels may be identified by numerous measurements. In this way, values

that were previously used but had lost their accuracy due to electrical changes over time-especially MPP values-can be recalculated and determined as real values once again. This recalibration ensures that the performance and efficiency of photovoltaic cells are accurately assessed, accounting for any variations or degradations that may have occurred over time.

It is advised to extract the I-V characteristics of solar panels independently, even if they have the identical manufacture features. Thus, the effect of even minor deviations will be reflected in the feature through measurement. It is hoped and suggested that this study will find a wide application area by demonstrating its usability in processes that require further study and application of I-V characteristic curves related to PV solar panels.

ETHICAL DECLARATIONS

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. NREL (National Renewable Energy Laboratory), "http://www.nrel.gov".
2. Akbaba, M. (2006). Optimum matching parameters of an MPPT unit used for a PVG-powered water pumping system for maximum power transfer. *Int J Energy Res*, 30(6), 395-409.
3. Brambilla, A., Gambarra, M., Garutti, A., & Ronchi, F. Optimum matching parameters of an MPPT unit used for a PVG- powered water pumping system for maximum power transfer. In proceedings of the 30th Power Electronics Specialists Conference, Charleston, USA, 1999, vol. 2, pp. 632-637.
4. Kulaksiz, A.A., & Akkaya, R. (2012). Training data optimization for ANNs using genetic algorithms to enhance MPPT efficiency of a stand-alone PV system. *Turk J Electr Engineer Comput Sci*, 20(2),241-254.
5. Ali, A., Almutairi, K., Padmanaban, S., Tirth, V., Algarni, S., Irshad, K., ... & Malik, M. Z. (2020). Investigation of MPPT techniques under uniform and non-uniform solar irradiation condition-a retrospection. *Ieee Access*, 8,127368-127392.
6. Motahhir, S., El Hammoumi, A., & El Ghzizal, A. (2020). The most used MPPT algorithms: review and the suitable low-cost embedded board for each algorithm. *J Clean Product*, 246,118983.
7. Yang, B., Zhu, T., Wang, J., Shu, H., Yu, T., Zhang, X., ... & Sun, L. (2020). Comprehensive overview of maximum power point tracking algorithms of PV systems under partial shading condition. *J Clean Product*, 268,121983.
8. Shankar, N., & Saravana Kumar, N. (2020). Reduced partial shading effect in multiple PV array configuration model using MPPT based enhanced particle swarm optimization technique. *Microprocessor Microsystems*, 103287.
9. Karami, N., Moubayed, N., & Outbib, R. (2017). General review and classification of different MPPT Techniques. *Renewabl Sustainabl Energy Rev*, 68,1-18.
10. Rezk, H., Fathy, A., & Abdelaziz, A.Y. (2017). A comparison of different global MPPT techniques based on meta-heuristic algorithms for photovoltaic system subjected to partial shading conditions. *Renewabl Sustainabl Energy Rev*, 74,377-386.
11. Hasanah, R.N., Setyawan, A.B., Maulana, E., Nurwati, T., & Taufik, T. (2020). Computer-based solar tracking system for PV energy yield improvement. *Int J Power Electr Drive Syst*, 11(2),743.
12. Zegrar, M., Benmessaoud, M.T., & Zerhouni, F.Z. (2021). Design and implementation of an IV curvetracer dedicated to characterize PV panels. *Int J Electr Comput Engineer*, 11(3),2011.
13. De Riso, M., Matacena, I., Guerriero, P., & Daliento, S. (2024). A wireless self-powered I-V curve tracer for on-line characterization of individual PV panels. *IEEE Transact Industr Electr*, 71(9),11508-11518.
14. Akbaba, M. (2007). Matching induction motors to PVG for maximum power transfer. *Desalination*, 209(1-3),31-38.
15. Türk, F. (2024). Investigation of machine learning algorithms on heart disease through dominant feature detection and feature selection. *Signal Imag Video Proces*, 18(4),3943-3955.
16. Turk, F. (2024). Rngu-net: a novel efficient approach in segmenting tuberculosis using chest X-Ray images. *Peer J Comput Sci*, 10,e1780.
17. Voutsinas, S., Karolidis, D., Voyiatzis, I., & Samarakou, M. (2023). Development of a machine-learning-based method for early fault detection in photovoltaic systems. *J Engineer App Sci*, 70(1),27.
18. González, I., Portalo, J.M., & Calderón, A.J. (2021). Configurable IoT open-source hardware and software IV curve tracer for photovoltaic generators. *Sensors*, 21(22),7650.
19. Sardar, R.H., Bera, A., Chattopadhyay, S., Ali, S.I., Pramanik, S., & Mandal, A.C. (2024). The impact of series (Rs) and shunt resistances (Rsh) on solar cell parameters to enhance the photovoltaic performance of f-PSCs. *Optical Materials*, 155,115818.
20. Hocine, L., Samira, K.M., Tarek, M., Salah, N., & Samia, K. (2021). Automatic detection of faults in a photovoltaic power plant based on the observation of degradation indicators. *Renewable Energy*, 164,603-617.
21. Oliva, D., Cuevas, E., & Pajares, G. (2014). Parameter identification of solar cells using artificial bee colony optimization. *Energy*, 72,93-102.
22. Appelbaum, J., & Sarma, M.S. (1989). The operation of permanent magnet DC motors powered by a common source of solar cells. *IEEE Transact Energy Convers*, 4(4),635-642.
23. Saied, M.M. (1988). Matching of DC motors to photovoltaic generators for maximum daily gross mechanical energy. *IEEE Transact Energy Convers*, 3(3),465-472.
24. López-Lapeña, O., Penella, M.T., & Gasulla, M. (2009). A new MPPT method for low-power solar energy harvesting. *IEEE Transact Industr Electr*, 57(9),3129-3138.
25. Akbaba, M., & Alattawi, M.A. (1995). A new model for I-V characteristic of solar cell generators and its applications. *Solar Energy Mater Solar Cells*, 37(2),123-132.
26. Goksenli, N., & Akbaba, M. (2016). Development of a new microcontroller based MPPT method for photovoltaic generators using Akbaba Model with implementation and simulation. *Solar Energy*, 136,622-628.

Wireless power transfer systems and wireless charging design between electric vehicles

Eda Bilgiç¹, Hasan Bilgiç², Emrah Kantaroğlu¹

¹Department of Mechanical Engineering, Faculty of Engineering, Kırıkkale University, Kırıkkale, Türkiye

²Department of Computer Engineering, Faculty of Engineering, Ahmet Yesevi University, Ankara, Türkiye

Cite this article: Bilgiç, E., Bilgiç, H., & Kantaroğlu, E. (2024) Wireless power transfer systems and wireless charging design between electric vehicles. *J Comp Electr Electron Eng Sci*, 2(2), 56-61.

Corresponding Author: Eda Bilgiç, edabilgic27@gmail.com

Received: 21/08/2024

Accepted: 07/10/2024

Published: 29/10/2024

ABSTRACT

The fundamentals of wireless power transmission began with Nikola Tesla, who used a Tesla coil to transmit energy wirelessly. These studies have paved the way for today's wireless power transmission technologies. In recent years, wireless power transmission has also garnered attention from phone companies and is now being considered for use in electric vehicles. One of the main reasons lithium-ion batteries are preferred in electric vehicles is the ease of charging. These batteries provide longer range compared to other types of batteries due to their high energy density, meaning they can store more energy relative to their size. Additionally, their faster charging capabilities are a major advantage for electric vehicle owners. As a result, charging technology is of great importance for electric vehicles. In this study, a design is presented that allows electric vehicles, which are becoming more common in daily life due to advancing technology, to wirelessly transfer charge to each other in emergency situations. In the design, the receiving coil is placed in the front hood of the vehicle, and the transmitting coil is placed in the rear hood, with the model being configured accordingly. The design also allows the location of the transmitting coil to be adjusted via distance sensors. When electric vehicles reach the optimum distance, the distance sensor gives an alert, allowing enough charge to be transferred for the vehicle to reach the nearest charging station. Additionally, the system can be designed to allow emergency vehicles (such as ambulances) to quickly receive energy from another vehicle in emergencies, due to the ease of sharing provided by the system. This system reduces the towing costs of electric and hybrid vehicles and offers drivers uninterrupted driving. Furthermore, the system provides numerous advantages such as eliminating cable clutter, offering ease of use, compatibility with all devices, multi-charging capability, and durability.

Keywords: Electric cars, wireless charging, transportation

INTRODUCTION

In recent years, the climate crisis and global warming have made the search for sustainable solutions even more critical. In response to the air pollution caused by fossil fuels, the importance of sustainable transportation has been steadily increasing. Under current conditions, electric vehicles have become one of the key components of transportation. Compared to internal combustion engine vehicles, electric cars offer an environmentally friendly and cost-effective energy alternative with a zero-emission principle. Based on the prediction that these electric vehicles will replace internal combustion engine vehicles worldwide and eventually dominate the market, the idea for this study emerged. In recent years, Türkiye has made remarkable strides, even envied globally, with the introduction of 100% electric TOGG vehicles and the establishment of a mass production facility in Gebze. Along with the production of 100% electric domestic and national vehicles, it is anticipated that vehicles with internal combustion engines currently in use will gradually be phased out. In other words, users

purchasing a brand-new electric vehicle will likely put their internal combustion engine cars up for sale (İpek, 2022). This indicates a major transformation, where the demand for electric vehicle charging stations is growing daily. However, the significant increase in the number of electric vehicles charging stations may still not be sufficient, as users will need to charge their vehicles at the end of long trips. Currently, limiting factors for the widespread adoption of electric vehicles include parameters such as ion structures, anion/cation material structures, normal charging times, and fast-charging deformations. Despite these challenges, the number of electric vehicles is rapidly increasing, and these issues are being addressed through new technologies. One of the alternative methods proposed to address the issues of battery charging times or fast-charging problems is improving charging infrastructures. In this study, a different model for charging infrastructure has been developed, exploring the possibility of wireless charging between vehicles. Relevant studies in literature have been examined.

Literature Review

Contactless electric vehicle charging systems are a research area that has attracted great attention in recent years to meet the energy needs of electric vehicles. These systems allow vehicles to be charged wirelessly while in motion or parked. Wireless charging, especially while on the move, can solve the range problem of vehicles and reduce dependence on charging stations by continuously transferring energy along the road. Most of the work in this field is based on inductive power transfer (IPT) technology. IPT provides energy transfer from the primary coil to the secondary coil via a magnetic field. In vehicle charging systems, this technology is used between a series of primary coils (transmitter) placed under the road and a secondary coil (receiver) placed under the vehicle. It is very important to determine the correct inductance values for the coils to work effectively. At this point, correct calculation of inductance is a critical factor in terms of energy transfer efficiency and system performance.

Studies in literature generally focus on the following areas.

Coil design and geometry: The inductance value depends on factors such as the size and shape of the coils, the number of turns, the distance between them and the use of magnetic cores. That's why coil design is important for energy efficiency.

Some studies provide optimized designs by analyzing how coil geometry affects inductance. For example, planar chokes and ferrite core chokes have been widely studied to improve efficiency.

Inductance calculation methods: There are various studies in which inductance is calculated using analytical, numerical and experimental methods. FEM (finite element method) is commonly used as a tool for these calculations.

Calculations based on electromagnetic principles such as Maxwell's equations and Ampere's law allow accurately model the magnetic field and inductance behavior of coils.

Energy transfer efficiency: In contactless charging systems, studies on resonance frequencies are important to reduce energy loss and increase efficiency. Inductive charging systems with magnetic resonance are used to optimize energy transfer.

Dynamic charging technologies: Charging the vehicle while it is in motion contributes to the expansion of the electric vehicle infrastructure by reducing the number of charging stations. The literature in this field is particularly focused on dynamic charging systems.

In addition to this literature survey, current advances in wireless charging systems along the way include new materials, electronic circuit designs, and more advanced control algorithms to increase the broad applicability of the systems.

With the increasing use of electronic devices, especially mobile devices, wireless charging methods have been developed to eliminate the disadvantages of wired charging systems. Thanks to wireless power transmission, people have become more mobile, and cable dependency has been removed, replacing the need for fixed positions and wiring. The simplicity and durability of wireless energy transfer offer much more effective solutions in such cases. Furthermore, wireless energy transfer is provided with high efficiency

(Karakaya, 2007). Wireless power transfer refers to the transmission of electrical energy from a power source to a target without any physical connection. The idea of wireless energy generation was introduced by Tesla at the beginning of the 1900s, and experimental studies were conducted to achieve this energy transfer. However, since the direction of the energy could not be controlled, the efficiency remained low (Arslan & Erkan, 2022). Today, despite the availability of different techniques, the most common wireless power transfer methods are magnetic resonance coupling, inductive coupling, and microwave power transfer techniques (Imura et al., 2009). In a study aimed at designing a wireless electric vehicle charging system, a wireless power transfer system with hexagonal coils of 1 kW was established, achieving an 85% charging efficiency performance in 10 cm of atmospheric air (Aydın & Aydemir, 2021). Kızıldağ and Yılmaz have researched and examined studies on wireless power transfer systems. They simulated the circuit in Matlab/Simulink and analyzed the advantages of power transfer methods during charging. At the end of their study, they concluded that inductive power transfer technology is efficient and reliable for transmitting electric power wirelessly (Kızıldağ & Yılmaz, 2021). In their studies, Yugendra Rao (2015), calculated the mutual inductance between a pair of coils for wireless power transfer applications using Ansys Maxwell, assigning the distance between the two coils as a variable. Tel & Kuşdoğan (2019), worked on charging electric vehicles using wireless power transmission with ansys maxwell simplorer. This study included the analysis results of the parameters on which the inductance between the coils depends, and the optimal coil selection was made.

In this study, based on the long-distance requirements of the vehicles, the design of the energy transfer between them for instantaneous vehicle charging is analytically designed. The study also investigates the possible ways to improve the energy efficiency of electron transfer during charging.

METHODS

In this study, the design of the coils that enable wireless power transfer is very important. To increase the efficiency of the system, the coils are designed by considering parameters such as material, thickness, length and distance of the terminals, number of turns, geometry and dimensions of the coils. In wireless energy transfer, whether the transmitter and receiver coils are spatially and angularly aligned also affects efficiency. If the coils are aligned, they can operate with higher efficiency in higher air gaps than angular and spatially unaligned coils (Ağçal & Doğan, 2021). When designing a wireless charger, features such as inductance, cross-section, number of windings and copper length should be considered in the selection of a suitable coil, and coil dimensions should also be considered. Reducing the size of the coils for the same number of turns decreases the inductance value of the coils. The size of the coils will therefore be optimally adjusted to suit all vehicles. The average air gap distance between the coils varies between 150-300 mm, in this study it is assumed that the distance difference is within this range. Once the design criteria have been determined, the prototype of the system has the potential to be applied to model vehicles.

In this study, the inductance value between the coils was first calculated analytically and then analyzed with Ansys

Maxwell. When designing a wireless charger, inductance, cross-section, number of windings and copper length should be considered in addition to coil dimensions. Reducing the coil dimensions for the same number of windings will decrease the inductance value of the coils. For this reason, the size of the coils is optimally adjusted to fit all vehicles.

The calculation of the circuit element values in the wireless power system is done with the following equations. In Equation 1, the values of the circuit elements, i.e. resistance elements, were calculated (Tang et al., 2012).

$$R = \frac{1}{S \cdot \sigma} \quad (1)$$

In Equation 2 and Equation 3, the values of system elements such as capacitance were calculated.

Primary Coil:

$$R_1 = 1.7 \times 10^{-8} * \frac{25}{49 \times 10^{-6}} = 0.0087 \Omega \quad (2)$$

Secondary Coil:

$$R_2 = 1.7 \times 10^{-8} * \frac{25}{49 \times 10^{-6}} = 0.0087 \Omega \quad (3)$$

In addition, the square coil cross-sectional area is 49 mm^2 , and the resistivity of copper is $1.7 \times 10^{-8} \Omega \text{m}$.

Inductance values of the designed system are given below.

For distance 60 mm:

LTx (L1)=119.86 μH , LRx(L2)=134.75 μH and $k=0.3781$.

The coil dimensions are determined as 400×400 , and the thickness of the wire is 7 mm in the system.

$$\omega_0 = 2\pi \cdot f_0 \text{ (angular frequency)} = 2\pi \cdot 10^3 = 62.83 \times 10^3 \quad (4)$$

$$C_p = \frac{1}{\omega_0^2 \cdot L_2} = \frac{1}{(62.83 \times 10^3)^2 \cdot (134.75 \times 10^{-6})} = 1.87 \mu\text{F} \quad (5)$$

$$C_s = \frac{1}{\omega_0^2 \cdot (1-k^2) \cdot L_1^2} = \frac{1}{(62.83 \times 10^3)^2 \cdot (1-0.3781^2) \cdot (119.86 \times 10^{-6})} = 2.3 \mu\text{F} \quad (6)$$

Finally, the calculations showed that the system is efficient when the source frequency is above 10 kHz.

According to this data, the efficiency was calculated as in Equation 7.

$$\text{Efficiency}(n) = \frac{P_{\text{out}}}{P_{\text{input}}} \times 100 = \frac{15.1250}{19.2129} \times 100 = 78.72\% \quad (7)$$

In the design, the voltage at the receiver coil is calculated as $V_{\text{out}} (\text{Rms}) = 194.4544 \text{ V}$ and the current value at the output is calculated as $I (\text{Rms}) = 77.7817 \text{ A}$ and according to these values, the receiver coil power is calculated in Equation 8. Power equation taken from (Croft, Terrell; Summers, Wilford I. 1987)

$$P_2 = I \cdot V = 77.7817 \cdot 194.4544 = 15124.99 \text{ W} = 15.125 \text{ kW} \quad (8)$$

Parameters Affecting the Analysis of Designed Coils

Analysis studies on the coils whose dimensions were determined were carried out on two different coil geometries: circular and square. According to the studies conducted by (Mahesh et al., 2021), it was determined that the effects of misalignment were less in square-sized coils and rectangular geometry was studied in the analysis. Coil sizes are determined by what the Maxwell student version can decipher. At the

prototype stage, size selection appropriate to reality will be made. Analysis was carried out in two different geometries and improvement efforts were directed according to the analysis results. In improvement works; Parameters such as coil geometry, number of windings, dimensions, materials used, structure of the core, and cross-section selection were studied. As a result of the analysis, the data was examined, and the most efficient coil distance was determined to be 50 mm in the selected dimensions. As the distance increases, the connection factor and inductance value decrease. For this purpose, the optimum distance should be selected in the preferred coil design. These parameters were considered in the study to reduce losses and increase system efficiency. The simulation of the selected coil structure of the design under specified conditions was solved in Ansys Maxwell and Ansys Maxwell simplorer programs.

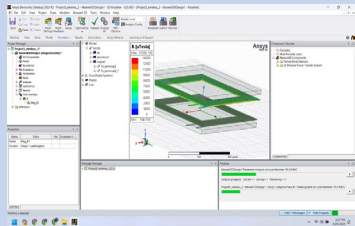
Characterization of Materials

Experiments were carried out on the materials of the coils: copper, steel 4300, steel 4340 and copper-coated silver. Analysis showed that copper has high electrical conductivity but low magnetic field. Steel was found to have a higher magnetic field and lower conductivity than copper. Material selection was based on criteria such as low cost, long life, high efficiency, conductivity and magnetic field generated. In recent years, efforts to increase efficiency on the coil and core have been ongoing. The analysis was performed at room temperature and in the air. As a result of the analysis, inductance values, coupling factor, efficiency and magnetic field value of the coils were calculated using ansys maxwell.

Analysis of Aluminum Core Coil in Square Dimensions

The properties of the sized aluminum core coil are shown in Table 1. Since aluminum is a paramagnetic material, it is a material with a relative permeability slightly greater than 1. When placed in a magnetic field, they are weakly magnetized in the direction of the field. Ferrite is a ferromagnetic material. They are substances with a relative magnetic permeability much greater than 1. When placed in a magnetic field, they magnetize very strongly in the direction of the field (Jackson & John; 1998 and Jiles & David; 1998). For this reason, the inductance value obtained from the coils we obtained using aluminum core was lower. Here material selection is a good way to increase efficiency.

Table 1. Analysis results of square aluminum core coil

Features of the transmitter coil		Features of the receiver coil	
Width=7 mm	NTx=14	Height=7 mm	NRx=14
Distance between windings =1 mm (2*polygon radius=radius change)			
Geometry of Square Aluminum core coil and core 400*400 mm			
Core thickness= 20 mm, Distance = 60 mm			
			
LTx =67.2204 μH		LRx = 72.6084 μH	
k=0.2357			

The analysis results of the aluminum core coil, the dimensions of which were determined, are shown in Figure 1. Looking at the results in Figure 1, the inductance value obtained for 70 mm is maximum. In addition, it was determined because of the analysis that the ferrite core coil used in the same dimensions is more efficient than the aluminum core coil.

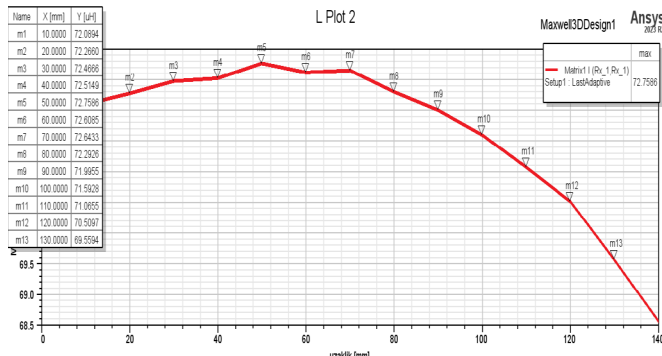
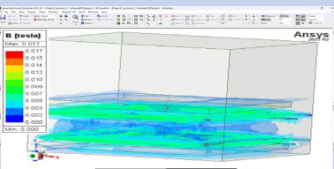


Figure 1. Change of inductance values in the receiver coil depending on distance

Analysis of Ferrite Core Coils with Square Dimensions

The properties of the ferrite core coil, the dimensions of which were determined, are shown in Table 2. The disadvantage of the ferrite core is that it has corrosion, and the way to prevent this is by coating the ferrite core. Or another alternative is to use an aluminum core over a ferrite core, which would be a design to reduce corrosion.

Table 2. Analysis results of the ferrite core coil in square dimensions

Features of the transmitter coil		Features of the receiver coil	
Widtht = 7 mm	NTx = 14	Height = 7 mm	NRx = 14
Distance between windings = 1 mm (2*poligon radius=radius change)			
Geometry of the quadratic ferrite core coil and core = 400*400 mm			
Core thickness 20 mm, Distance = 60 mm, Magnetic field = 0.017 T			
			
LTx = 119.8622 μH		LRx = 134.7552 μH	
k = 0.3781		M = 47.874 μH	

The analysis results of the ferrite core coil, whose dimensions were determined, are shown in Figure 2, 3. As can be seen from the analysis, when ferrite material was used as the core, inductance values increased more than aluminum. The reason for this is that ferrite material has a better magnetic permeability than aluminum material.

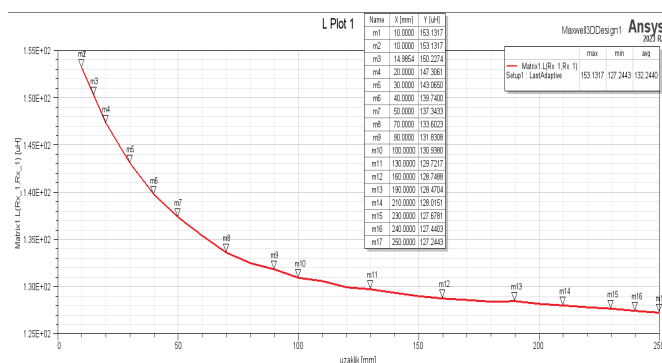


Figure 2. Change of inductance values in the receiver coil depending on distance for the ferrite core coil

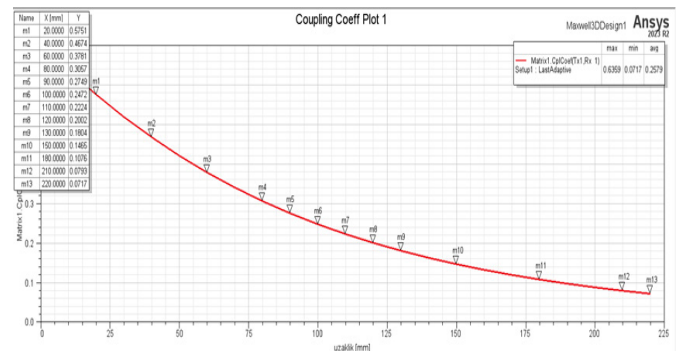


Figure 3. Change of coupling factor in coils depending on distance for the ferrite core coil

RESULTS

In this study, a design is presented that enables electric vehicles, which are becoming more and more widespread in daily life with the developing technology, to transfer their loads to each other in emergency situations in a contactless way. The 3D models are transferred to simplorer file and the efficiency of the system is calculated. The simplified structure of the system will be prepared in Ansys Maxwell simplorer and the circuit structure of the system is prepared by making numerical calculations according to the inductance values obtained from the coils. As a result of the analysis, a sinzoidal graph was obtained and the efficiency of the system in the charging state was found to be 78.72%. Simplorer circuit results for these values are as follows. The geometry we created in ansys maxwell 3D and transferred to the simplified wireless power transmission system is shown in Figure 4.

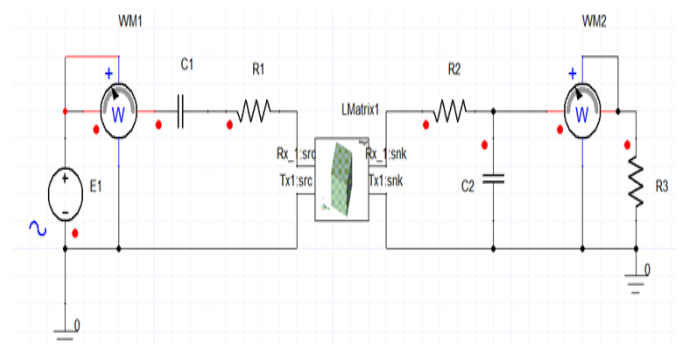


Figure 4. Importing Ansys Maxwell 3D component to simplorer file

The simulation results obtained by transferring the geometry created in Ansys Maxwell 3D to the simplified wireless power transmission system are shown in Figure 5.

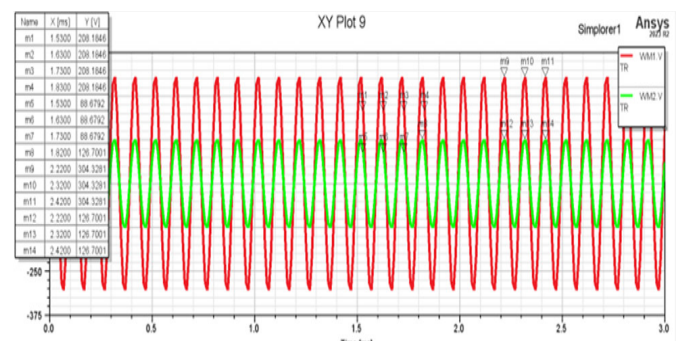


Figure 5. Graph of the sinusoidal input and output voltage resulting from the analysis

The resonant circuit shown in Figure 6 is made to avoid losses that may occur during the transfer of the Maxwell file. The result of the analysis of the resonant circuit model is shown in

Figure 7. The simplified resonant circuit model is taken from (Aditya, 2016).

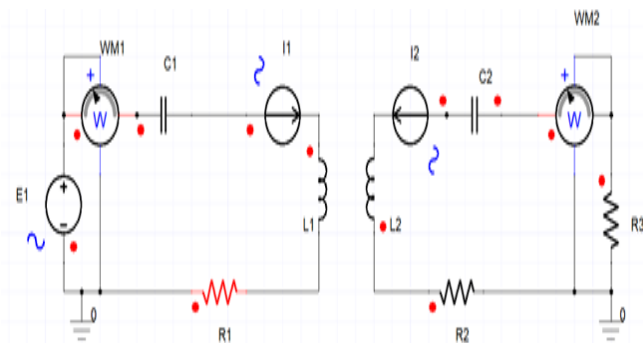


Figure 6. Creating the resonance circuit model in simplorer

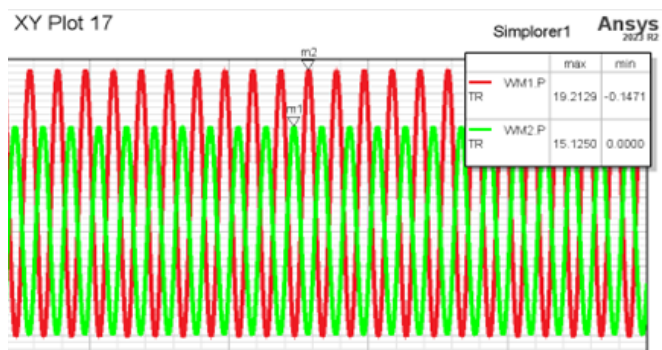


Figure 7. Graph of input and output power as a result of the analysis of the resonance circuit model

Limitations

The analyzes used in the study could not be supported by experimental studies. The system has limitations such as charging speed, heating problems, distance restriction and misalignment. These limitations may decrease as studies progress. Studies that can be done in the future; the expected performance can be achieved with studies such as advanced cooling systems, moving alignment and material improvements.

CONCLUSION

In short, in this study; vehicles will be able to charge enough to reach the charging station via wireless power transmission. In this way, drivers will be able to control the process in emergency situations without panicking. When designing the wireless charger, inductance, cross-section, number of windings, copper length, as well as coil dimensions should be taken into consideration when selecting the appropriate coil. Decreasing the coil dimensions for the same number of windings reduces the inductance value of the coils. Therefore, the size of the coils will be optimally adjusted to be compatible with all vehicles. The magnetic field formed in ferrite-core copper coils is below 0.02 Tesla, meaning that the system alone is not at a level that can affect human health. ICNIRP (ICNIRP, 1998) and IEEE standards were taken as reference to examine the harms of the wireless power transmission system to human health IEEE, (2019). The magnetic field emitted by the system is below these reference values. However, we continue to work on the magnetic field that will occur if more than one user activates the system in heavy traffic. Protective shield works to protect against magnetic fields will be concentrated in the future. If we compare it with the article by (Tel & Kuşdoğan; 2019), trial studies

have been carried out on different materials. It can reduce the disadvantage of other materials, which causes an increase in the power drawn from the source due to their low conductivity, by charging the battery with renewable energy or by coating the coils. This study has the following limitations.

ETHICAL DECLARATIONS

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Acknowledgments

We would like to thank the Ministry of Transport and Infrastructure for supporting this work with the “Reaching and Accessing Türkiye 2053 Idea Contest”.

REFERENCES

- Aditya, K. (2016). Design and implementation of an inductive power transfer system for wireless charging of future electric transportation. A Thesis, University of Ontario Institute of Technology Oshawa, Canada, 1-190.
- Ağçal, A., & Doğan, T.H. (2021). 1 Kw gücündeki enerji transfer sistemi tasarımı ve insan sağlığına etkileri. *Müh Bil Tasar Derg*, 9(3),1-10.
- Arsalan, M.T., & Erkan, K. (2022). Literature review on wireless power transmission by magnetic resonance coupling method. *Eur J Sci Technol*, 42,118-125. doi:10.31590/ejosat.1187350
- Aydın, E., & Aydemir, M.T. (2021). A 1-kW wireless power transfer system for electric vehicle charging with hexagonal flat spiral coil. *Turk J Electric Eng Comput Sci*, 29(5),2346-2361. doi:10.3906/elk-2012-68
- Croft, T. & Summers, W. (1987). American electricians' handbook (Eleventh ed.). New York: McGraw Hill. ISBN 0-07-013932-6.
- IEEE, (2019). Standard for safety levels with respect to human exposure to electric, magnetic, and electromagnetic fields, 0 Hz to 300 GHz. IEEE Std C95.1-2019. 1-312. doi:10.1109/IEEESTD.2019.8859679
- Imura, T., Okabe, H., & Hori, Y. (2009). Basic experimental study on helical antennas of wireless power transfer for electric vehicles by using magnetic resonant couplings. 5th IEEE Vehicle Power and Propulsion Conference, VPPC '09, 936-940.
- International Commission on Non-Ionizing Radiation Protection (ICNIRP), (1998). Guidelines for limiting exposure to time varying electric and magnetic fields for low frequencies (1 Hz–100 kHz). *Health Phys*, 99,818-836.
- İpek, B. (2022). Electric vehicles, electric charging stations and electric vehicle quantity future forecast for Turkey (Dissertation). İstanbul Technical University, Institute of Science and Technology, Türkiye.
- Jackson, J.D. (1998). Classical electrodynamics (3rd ed.). New York: Wiley. p. 193. ISBN 978-0-471-30932-1.
- Jiles, D. (1998). Introduction to magnetism and magnetic materials. CRC Press, 354. ISBN 978-0-412-79860-3.
- Karakaya, U. (2007). Motor control via wireless energy and information transfer (Dissertation). İstanbul Technical University, Institute of Science and Technology, Türkiye.

13. Kızıldağ, U. & Yılmaz, A.S. (2021). Wireless charging systems in electric vehicles and a sample system investigation. *J Eng Sci Adiyaman Uni*, 8(14), 209-224.
14. Mahesh, A., Chokkalingam, B., & Mihet-Popa, L. (2021). Inductive wireless power transfer charging for electric vehicles-a review. *IEEE Access*, 9,137667-137713. doi:10.1109/ACCESS.2021.3116678
15. Tang, Z., Cristini, M., & Koga, T. (2012). Wireless power transfer using maxwell and simplorer. *Automotive Simulation World Congress*.
16. Tel, Ö.F., & Kuşdoğan, Ş. (2019). Charging of electric vehicles with wireless power transmission. *Kocaeli Üni Fen Bil Derg*, 2(1),15-26.
17. Yugendra Rao, K.N. (2015). Dynamic modeling and calculation of self and mutual inductance between a pair of coils for wireless power transfer applications using ANSYS Maxwell. *Int Adv Res J Sci Eng Technol*, 2(10),1-3. doi:10.17148/IARJSET.2015.21002

Vision assistant for visually impaired individuals

İrem Yıldırım, İlyas Demir, İsa Peker, Mehmet Emin Yılmaz, Ozan Güneyli,
Ezgi Şevval Aktaş, Elif Özer, Emrah Kantaroğlu

Department of Electronics and Automation, Faculty of Engineering and Natural Sciences, Kırıkkale University, Kırıkkale, Türkiye

Cite this article: Yıldırım, İ., Demir, İ., Peker, İ., Yılmaz, M.E., Güneyli, O., Aktaş, E.Ş., Özer, E., & Kantaroğlu, E. (2024) Vision assistant for visually impaired individuals. *J Comp Electr Electron Eng Sci*, 2(2), 62-66.

Corresponding Author: İrem Yıldırım, iremyldrm5706@gmail.com

Received: 25/09/2024

Accepted: 18/10/2024

Published: 29/10/2024

ABSTRACT

Visually impaired individuals face significant challenges in many areas of their daily lives, such as accessing environmental information, social interaction, and safety. This article aims to enhance the independence of visually impaired individuals and facilitate their daily lives through an artificial intelligence-supported image and sound recognition system. The developed prototype consists of glasses equipped with a Raspberry Pi 4 microprocessor, integrated headphones, touch controls, and a robust battery management system. Its electronic schematic includes the integration of the camera and the microprocessor. The artificial intelligence algorithms developed successfully perform tasks such as object recognition, text reading, and face recognition. A visual question answering (VQA) model was utilized in the glasses to enable visually impaired individuals to communicate more effectively with their surroundings. Following a demonstration conducted with visually impaired individuals, the usability and effectiveness of the prototype were evaluated, and improvements were made based on user feedback. The accuracy of the algorithms was tested, resulting in a prototype with a 78% accuracy score, which was close to the 82% accuracy score of GPT-4. Additionally, tests conducted with 20 users indicated an 85% user satisfaction rate.

Keywords: Glasses, artificial intelligence, image processing, mini computer, visually impaired individuals, 3D CAD

INTRODUCTION

According to the World Health Organization (WHO), approximately 45 million people globally are completely blind, and 135 million have partial vision loss (WHO, 2003). As of 2022, the Turkish Ministry of Family and Social Services reported that there are 1,039,000 visually impaired individuals in Turkey, with 478,000 men and 561,000 women (Ministry of Family and Social Services, 2022). Common problems for both fully and partially visually impaired individuals include independence, education, transportation, social life, and societal pressure.

Visually impaired individuals often rely on healthy individuals until they can handle the challenges they face throughout life. This reliance creates problems for both the visually impaired person and the helper. Moreover, a helper may not always be available. Visually impaired individuals use the Braille alphabet in the field of education. However, the limited number of resources available in Braille and the lack of educators who can teach using this alphabet make accessing information a challenge.

Transportation is another problem for visually impaired people. The lack of sufficient tactile guide paths in public

spaces, the absence of guiding structures within buildings, and obstacles placed on sidewalks by unaware individuals prevent visually impaired people from moving safely. Additionally, societal exclusion and lack of participation in social events negatively impact their psychological health, causing some to withdraw from social life.

As a solution to these problems, we designed an AI-powered vision assistant in the form of glasses. The user can take photos with the glasses using a button on the side, and the camera captures visual information quickly and in high resolution. The camera system provides a wide field of view and allows detailed image processing. The photo is analyzed by AI using an image processing algorithm, which includes both image recognition and scene understanding. The system recognizes objects, reads text, and identifies faces and environments, then communicates this information to the user via headphones. The speed, tone, and language of the descriptive sounds can be adjusted according to user preference. The glasses feature a user-friendly interface with touch and voice control systems, ensuring accessibility and comfort during prolonged use.

The system is designed with privacy in mind, using encryption to protect user data and providing offline data processing options. This ensures that user data is processed securely without reliance on cloud-based systems, maintaining privacy.

LITERATURE REVIEW

Several studies have been conducted on smart glasses for visually impaired individuals. These studies examine the benefits of such technologies in enhancing daily life and promoting independence. In the literature, four prominent studies highlight how different approaches and technologies assist visually impaired people. These studies address the functions of smart glasses, such as facial recognition, object detection, and navigation.

Bastola et al, focused on the development of multifunctional glasses for visually impaired individuals, allowing users to receive auditory and tactile notifications about objects, people, and environmental obstacles. The glasses feature text recognition technology for reading signs and written materials, and navigation tools to help users find their way in unfamiliar environments. Despite the advanced features, high cost and battery life issues are identified as disadvantages (Bastola et al., 2023).

Nazim et al, developed smart glasses using computer vision technologies. Their glasses use deep learning methods for object recognition, facial recognition, and Optical Character Recognition (OCR). Users can activate the glasses with simple commands and receive detailed descriptions of their surroundings. The glasses offer high accuracy but suffer from high power consumption and environmental sensitivity (Nazim & Firdous, 2023).

Mustafa et al, integrated various models into smart glasses designed for visually impaired individuals. These glasses provide features such as facial recognition, object detection, and navigation, powered by Raspberry Pi 4. The glasses offer low-cost solutions and are equipped with vibration alerts for obstacle avoidance. However, technical expertise is required for maintenance, and vibration alerts may discomfort the user (Mustafa, Omer, & Mohammed, 2023).

Waisberg et al, evaluated the potential benefits of second-generation smart glasses developed by Meta in collaboration with Ray-Ban for visually impaired individuals. These glasses feature a 12-megapixel wide-angle camera and advanced audio recording, aimed at improving the independence of visually impaired individuals in daily life. Although the glasses offer high accuracy and advanced technology, high cost and battery life issues are identified as challenges (Waisberg et al., 2023).

Visual data processing and machine learning play a significant role in various fields, such as medical diagnosis processes and interactive systems developed for visually impaired individuals. In particular, deep learning-based models used in medical imaging provide critical data for the accurate detection and diagnosis of diseases.

The processing and interpretation of visual data are applied in various fields, similar to their use in tuberculosis diagnosis. For example, deep learning-based models like U-NET provide effective segmentation in medical imaging

data, assisting experts in their decision-making processes. Similarly, VQA-based models can offer auditory feedback for visually impaired individuals, allowing them to have a more interactive experience in their environment (Türk, 2024).

The importance of machine learning algorithms in medical diagnosis is increasingly growing, and their performance significantly improves when combined with optimization methods. In a study conducted by Türk, the accuracy of machine learning models in classifying heart disease was increased to 99.08% using the dominant feature detection (DFD) method. Likewise, VQA-based models can help visually impaired individuals perceive and identify objects in their surroundings, enabling them to gain more independence in their daily lives (Türk, 2024).

This article presents a user-friendly design, low cost, long battery life, durability against environmental factors, Turkish language support, a modular structure, and easy integration, distinguishing it from other glasses. Instead of complex technological structures, we accelerated users' adaptation process with an intuitive and easy-to-use interface. By utilizing lower-cost components, we increased accessibility and facilitated the access of visually impaired individuals to these technologies. We extended battery life with advanced battery management systems and energy-efficient components. We developed systems that are minimally affected by environmental factors such as light, noise, and weather conditions. We provided a modular structure that allows for easy replacement and upgrading of system components.

METHODOLOGY

This smart glasses design, developed based on the needs of visually impaired individuals, stands out with its user-friendly and ergonomic structure. The prototype includes hardware such as a camera system, a microprocessor (Raspberry Pi 4), integrated headphones, and touch control sensors. Figure 1 shows the electronic components on the glasses.

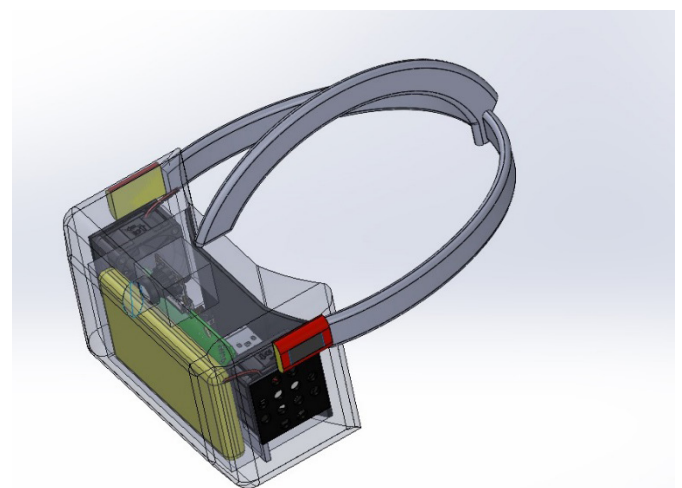


Figure 1. The electronic hardware of the glasses

To provide clear and high-quality images, the optical image stabilization camera integrated into the upper center of the glasses can capture sharp images even during movement, thanks to its dual-axis stabilization technology. This camera, with a resolution of 5 MP, has a 70° field of view, allowing it to cover a wide area.

The integrated Raspberry Pi 4 in the glasses serves as a mini-computer with its quad-core ARM Cortex-A72 processor. This processor can run image recognition and description software, providing users with auditory feedback. Additionally, it has the capability to respond to user questions within the images. The Raspberry Pi 4 is equipped with modern features such as 8 GB RAM, a MicroSD card slot, USB-C power input (5V, 3A), Wi-Fi, Bluetooth, and USB ports.

The integrated headphones in the glasses allow for the auditory transmission of images processed by artificial intelligence. With a frequency range of 20,000 Hz, sounds can be clearly heard even in the noisiest environments. The headphones come with a 1.2-meter long cable and a 3.5 mm jack input, providing a widely used connection standard worldwide.

The device's power requirement is met by a 7.4 V 2S 2200 mAh 35C LiPo battery. To prevent overheating of this battery, which supplies the necessary power to all electronic systems, cooling is provided by two mini fans integrated into the ventilation holes located at the back of the glasses. The fans, measuring 40x40x20 mm, operate at 5V and offer effective cooling performance at a noise level of 30.5 dBA. This ensures the safe use of the battery.

The camera, headphones, and touch buttons are connected to the Raspberry Pi 4. The energy management of the glasses is handled by a regulator board, which reduces the voltage supplied by the battery to appropriate levels, ensuring stable operation of the electronic components. The regulator board operates at a voltage range of 1.25V to 30V.

Under the cover located on the left side, there is a processor that runs the software for recognizing objects and providing auditory descriptions. This processor processes images from the 2K resolution camera positioned in the center of the glasses and presents information to the user. Users can take photos using a button located behind the processor, and these images are audibly described by artificial intelligence. The described images are transmitted to the Bluetooth-connected headphones located at the back left of the glasses.

On the right side of the glasses, there is a battery that supplies power to all hardware. To prevent the battery from overheating, the cooling fan activates, effectively distributing heat through the ventilation holes on the glasses.

The prototypes are designed as innovative smart glasses aimed at facilitating the daily lives of visually impaired individuals. They combine high performance and ergonomics with modern hardware components and a user-friendly interface. The glasses are designed using a 3D CAD program. The prototypes of the glasses have been printed using PLA material with a 3D printer. Figures 2 and 3 showcase two different glasses designs for completely visually impaired individuals, while Figures 4 and 5 display designs for partially visually impaired individuals.

This article utilizes a visual question answering (VQA) based model that enables visually impaired individuals to communicate more interactively with their surroundings.

The operational principles and functionality of the VQA-based model have been described. The system is activated by pressing a button by the user and provides instant responses

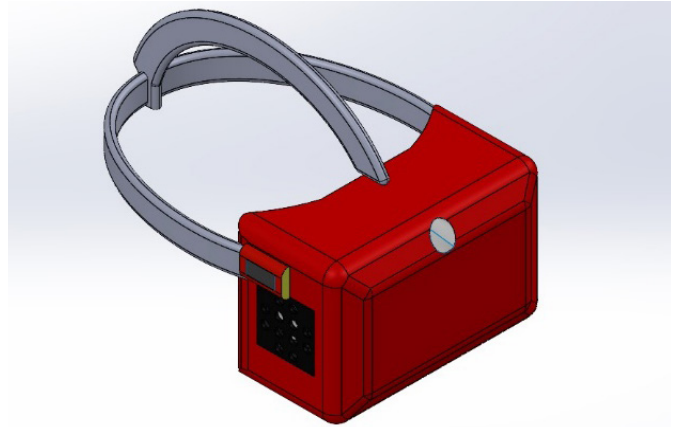


Figure 2. 3D CAD drawing of the glasses designed for completely visually impaired individuals



Figure 3. The prototype of the glasses designed for completely visually impaired individuals



Figure 4. 3D CAD drawing of the glasses designed for partially visually impaired individuals



Figure 5. The prototype of the glasses designed for partially visually impaired individuals

about environmental visuals through the individual's spoken questions. The working principles of the model are illustrated in Figure 6, and its functionality is explained step by step in the flowchart provided in Figure 7.

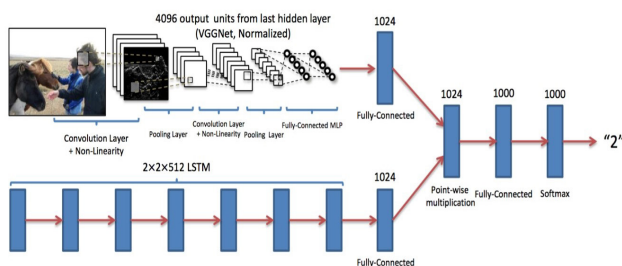


Figure 6. Visual question answering architecture (Aishwarya Agrawal et al., 2016)

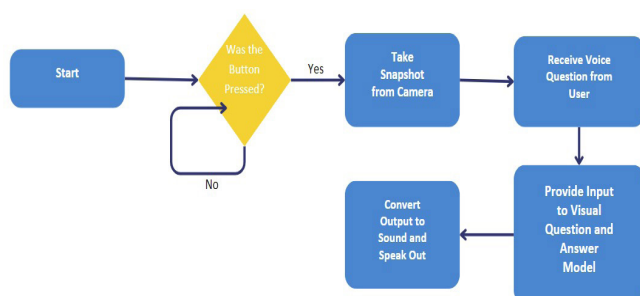


Figure 7. VQA-based visual question-answering model

This image is used as an example showing how the visual question answering (VQA) process works.

The VQA model answers the user's image-based questions by combining both visual and linguistic inputs. The key components of this process are as follows:

Image processing [Convolutional neural network (CNN)]: When an image is provided to the VQA model, it passes through a series of convolutional layers that extract meaningful features from the image. These layers analyze important details such as the edges, textures, and structures of the image. As a result, a normalized feature vector is obtained to be transferred to the fully connected layers.

Natural language processing [Recurrent neural network/long short-term memory (RNN/LSTM)]: Simultaneously, the question posed by the user is transferred to the model's language processing section, namely the LSTM layers. LSTM processes the words of the question sequentially and learns the meaning of the question, thus extracting linguistic features.

Feature fusion: The features extracted from the image are combined with the linguistic features derived from the question. This process is typically performed using the dot product method. The fusion enables the meaningful integration of important information from both the image and the question.

Answer generation: The combined features are sent to a classifier (fully connected layer), where a probability distribution is generated among possible answers through the softmax function. The model selects the answer with the highest probability, and this answer is presented to the user.

This article utilizes a visual question answering (VQA) based model that enables visually impaired individuals to

communicate with their surroundings in a more interactive way. The system becomes active when the user presses a button and provides instant responses about environmental visuals through the individual's spoken questions. The operation of the flowchart provided in Figure 6 is explained step by step below:

Initial state: The system is in an initial state and remains in passive mode until an input is received from the user.

Has the button been pressed?: The system checks whether the user has pressed the button.

- **Yes:** If the button has been pressed, the system moves to the next step.
- **No:** If the button has not been pressed, the system continues to remain in passive mode.

Capture instant image from the camera: After the user presses the button, the system captures an image of the current environment through the connected camera. This step ensures that the visual required to answer the user's question is obtained.

Receive voice question from the user: Once the visual is obtained, the system captures the user's question in a spoken format. This question can relate to any object, event, or situation in the visual.

Send visual and question to the VQA model: The system sends the voice question received from the user and the instant image obtained from the camera as input to the VQA model. The VQA model generates the correct answer based on the given question and image.

Convert output to speech and vocalize: The written response from the VQA model is converted into an audible form for the visually impaired individual and conveyed to the user through a speaker. This step allows the user to receive a direct and quick response to their question.

This system helps visually impaired individuals perceive their surroundings more independently and comfortably. By combining the instant images obtained from the camera with the spoken questions from the user, the system provides instant feedback. This system, which processes visual and auditory inputs to generate spoken responses, demonstrates the significant benefits of visual question answering (VQA) technology for visually impaired individuals.

The system employs a unique image processing algorithm that offers both image recognition and scene understanding capabilities. This advanced algorithm provides high accuracy and speed in tasks such as object recognition, text reading, facial recognition, and environmental description. With more accurate descriptions and fast response times, the user experience is significantly improved.

In addition to the standard datasets in the literature, new datasets have been created considering specific situations that visually impaired individuals may encounter. These datasets have been optimized to be most suitable for the needs of visually impaired users.

The speed, tone, and language of the descriptive voices can be adjusted according to the user's preferences. The ability for users to personalize the audio description options is a rare feature in existing products. This customization option makes the system more accessible to users with different language and tone preferences.

The glasses feature a user-friendly interface that allows for easy interaction with touch and voice control systems. With its ergonomic design, it provides comfort during prolonged use and enhances accessibility.

The system features advanced security protocols that provide encryption and offline data processing options to secure user data. This privacy-focused approach allows data to be processed locally without relying on cloud-based systems, thereby maximizing personal privacy. Thanks to the offline operation capability, the system can process user data (images and voice questions) without needing an internet connection and does not export any data. By ensuring that data is not sent to any cloud service or third-party server, potential data breaches are prevented, keeping user information completely confidential. This structure offers a significant security advantage, ensuring that the data remains safe.

DISCUSSION

The design of the glasses aligns with existing habits, providing users with ease of use. The rapid description of images captured by the cameras enables users to receive immediate information. The glasses' form factor makes them lightweight, portable, and easy to wear. With camera glasses, information can be gathered discreetly and without drawing attention to the user. As the independent mobility of visually impaired individuals increases, their quality of life improves. These glasses can be used across a wide range of daily activities, from shopping to social interactions.

However, high processing power and continuous camera use can increase battery consumption, leading to limited usage time. The use of advanced technology and components may raise costs, making the product expensive for some users. Ensuring user privacy and the security of collected data can be a significant challenge. Integrating image processing and audio description systems can be complex and may lead to technical issues.

RESULTS

We measured the accuracy and speed of object recognition, text reading, and face recognition algorithms in the glasses. The prototype was introduced to visually impaired individuals, and the ease of use and effectiveness of the prototype were evaluated. Improvements were made based on user feedback.

CONCLUSION

When compared to one of the most popular models today, GPT-4, our model achieved results that are quite close, despite having only 500 million parameters compared to GPT-4's 1.76 trillion parameters. In tests conducted on 100 samples, GPT-4 achieved an accuracy score of 82%, while our model achieved a score of 78%.

ETHICAL DECLARATIONS

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

The authors declared that this study has received no financial support.

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Bastola, A., Enam, M. A., Bastola, A., Gluck, A., & Brinkley, J. (2023). Multi functional glasses for the blind and visually impaired: design and development. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. doi:10.1177/21695067231192450.
2. Nazim, S., Firdous, S., Shukla, V. K., & Pillai, S. R. (2023). Smart glasses: a visual assistant for the blind. Amity University Dubai.
3. Mustafa, A., Omer, A., & Mohammed, O. (2023). Intelligent Glasses for Visually Impaired People. Sudan University for Science and Technology.
4. Waisberg, E., Ong, J., Masalkhi, M., Zaman, N., Sarker, P., Lee, A.G., & Tavakkoli, A. (2023). Meta smart glasses-large language models and the future for assistive glasses for individuals with vision impairments. *Eye*, 38,1036-1038. doi:10.1038/s41433-023-02842-z
5. Turk, F. RINGU-NET: a novel efficient approach in segmenting tuberculosis using chest X-ray images. *PeerJ Comp Sci*, 2024,5,10. doi: 10.7717/peerj-cs.1780
6. Türk, Fuat. (2024). Investigation of machine learning algorithms on heart disease through dominant feature detection and feature selection. *Signal Image Video Process*, 18,1-13. doi:10.1007/s11760-024-03060-0
7. Agrawal, A., Lu, J., Antol, S., Mitchell, M., Zitnick, C. L., Batra, D., & Parikh, D. (2016). VQA: Visual question answering. *arXiv*, <https://arxiv.org/abs/1505.00468>.
8. Papers with Code. (n.d.). Visual Question Answering. Retrieved from <https://paperswithcode.com/task/visual-question-answering-1>
9. OrCam MyEye 2 Pro. (n.d.). Retrieved from <https://www.orcam.com/en-us/orcam-myeye-2-pro>
10. Envision Glasses. (n.d.). Retrieved from <https://www.letsenvision.com/glasses/home>
11. Berry, T. (2018, February 9). Visually stunning; aira smart glasses help the blind see airport gates. CW39. Retrieved from <https://cw39.com/2018/02/09/visually-stunning-aira-smart-glasses-help-the-blind-see-airport-gates>.
12. Wikipedia contributors. (2023, December 9). Seeing AI. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Seeing_AI.