iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XII Month of publication: December 2025

DOI: https://doi.org/10.22214/ijraset.2025.76066

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

Brain Tumor Detection Using SVM and Machine Learning Model

Parineeta Chundawat
Department of CSE, IGDTUW, Delhi, India

Abstract: The project focuses on detecting brain tumors using a Support Vector Machine (SVM) model combined with machine learning techniques to enhance diagnostic accuracy. Medical imaging data, specifically MRI scans, are processed and analyzed to identify tumor presence by classifying brain tissues as either normal or abnormal. Preprocessing steps, including noise reduction and segmentation, are applied to improve image clarity and focus on critical regions. The SVM model is trained on labeled datasets, allowing it to detect tumor patterns efficiently. This approach aims to aid early diagnosis, enhancing treatment outcomes by providing a reliable, automated solution for tumor identification.

The early detection of brain tumors remains one of the most demanding tasks in medical imaging, as even minor structural abnormalities in the brain can significantly impact diagnosis and treatment outcomes. This research presents a machine learning–based system designed to automatically identify brain tumors from MRI scans, with a primary focus on the Support Vector Machine (SVM) classifier. The approach begins with extensive preprocessing of MRI images, where noise removal, contrast enhancement, and region-of-interest segmentation are applied to highlight critical features while reducing irrelevant visual information. These refined images allow the system to extract meaningful patterns that distinguish healthy brain tissue from tumor-affected regions. The SVM model is trained on a labeled dataset of MRI scans and optimized to handle the complex, high-dimensional nature of medical imaging data.Its strength in binary classification makes it particularly suitable for identifying whether a tumor is present. For comparative analysis, additional models such as K-Nearest Neighbors (KNN) and regression techniques are also implemented. Experimental results show that SVM consistently achieves superior accuracy and produces more reliable predictions than the other models evaluated in this study. By integrating machine learning algorithms with systematic preprocessing techniques, the proposed framework demonstrates its potential as a supportive diagnostic tool for radiologists. The system not only accelerates the detection process but also reduces the risk of human error, making it highly beneficial for clinical settings where early diagnosis is crucial. Overall, this study highlights the effectiveness of SVM-based classification in brain tumor detection and lays the foundation for future enhancements such as tumor grading, treatment recommendation modules, and advanced analytical reporting.

Index Terms: SVM, ML Model, Jupiter Notebooks, Skit Learn, Python, NumPy, Pandas, OpenCV.

I. INTRODUCTION

Brain tumors are one of the most challenging medicals conditions due to their complex nature and critical location within the brain. Early and accurate diagnosis is crucial for effective treatment and improved patient survival rates [5]. Conventional diagnostic techniques, though effective, are time- consuming and often require expert interpretation, making automated detection methods highly desirable [6]. Machine learning (ML) has emerged as a powerful tool for medical diagnostics, offering the potential to analyze complex datasets and recognize subtle patterns that may elude human observation. [7]. In the research, a Support Vector Machine (SVM) model is employed for brain tumor detection using MRI scan data. SVM, a supervised ML algorithm, is particularly effective for binary classification. Tasks, making it an ideal choice for distinguishing between normal and abnormal brain tissues [8][9]. Brain tumors represent one of the most complex and life-threatening medical conditions, primarily due to their location within the central nervous system and the difficulty of identifying early-stage abnormalities. Even small tumors can disrupt essential neurological functions, making timely and accurate diagnosis extremely important for effective treatment and improved patient outcomes. Traditionally, radiologists rely on manual analysis of Magnetic Resonance Imaging (MRI) scans, a process that interpretation. However, the increasing number of requires extensive expertise and careful diagnostic imaging cases and the subtlety of early tumor patterns make manual assessment both time-consuming and error prone.

With the rapid advancement of computation in the medical field, machine learning has emerged as a powerful tool for supporting medical diagnosis. Machine learning techniques can examine large volumes of imaging data, identify hidden patterns, and detect abnormalities that may not be immediately visible to the human eye. Supervised machine learning algorithms have shown significant promise in image classification tasks, where the objective is to differentiate between normal and abnormal structures. Among these algorithms, the Support Vector Machine (SVM) model has gained considerable attention due to its reliability in handling high dimensional data and its strong performance in binary classification problems. In the context of brain tumor detection, MRI scans offer high-quality structural information that is suitable for automated analysis. However, raw MRI images often contain unwanted noise, intensity variations, and complex textures that can affect classification accuracy. Therefore, preprocessing techniques such as noise removal, normalization, and segmentation play an essential role in ensuring that the machine learning model receives clean and meaningful input data. Once the MRI images are pre-processed, feature extraction methods help identify crucial imaging characteristics that represent the presence of a tumor. This research focuses on developing an automated system for brain tumor detection using an SVM classifier, supported by additional models such as K- Nearest Neighbors (KNN) and regression techniques for performance comparison. By leveraging Python libraries such as Scikit-learn, NumPy, Pandas, and OpenCV, the study provides an efficient and structured workflow for training, validating, and evaluating the machine learning models. The aim of this research is not only to achieve accurate tumor detection but also to propose a method that can assist medical professionals by reducing diagnostic burden, improving decision- making, and supporting early intervention. Overall, the introduction of machine learning into medical imaging has created new possibilities for enhancing diagnostic accuracy and minimizing human error. Through this study, we demonstrate how SVM- based classification can contribute to the development of reliable, automated diagnostic tools that can be integrated into real-world clinical environments for improved patient care.

II. LITERATURE REVIEW

Brain tumor detection using medical imaging has been a significant area of research, with machine learning models, particularly Support Vector Machines (SVMs), emerging as reliable tools for this purpose. SVMs have demonstrated high accuracy in classifying brain tumor images due to their ability to handle nonlinear separations and provide robust decision boundaries, as highlighted by Cortes et al., 1995 [8]. Further research by Zhang et al., 2019 [3] demonstrates the applicability of SVMs in medical imaging, emphasizing their role in identifying tumors with precision.

The importance of preprocessing medical images before analysis is extensively discussed in studies such as Gupta et al., 2021 [2], which highlight how improved preprocessing techniques enhance tumor detection accuracy in MRI scans. Similarly, Soni et al., 2014 [11] emphasize preprocessing steps like noise reduction and normalization to prepare data for efficient machine learning model applications. These methods ensure that models like SVMs are trained on high-quality input data, which is critical for accurate predictions.

Recent advancements in combining SVMs with other machine learning algorithms have further improved tumor classification accuracy. For instance, Li et al., 2020 [4] demonstrated that automated classification systems leveraging ensemble methods and SVMs can outperform traditional diagnostic techniques. Kumar et al., 2018

[9] also emphasize the use of SVMs for medical image classification, citing their ability to work effectively in scenarios with limited training data—a common challenge in medical imaging. Moreover, the role of feature selection in SVM-based models is crucial. Studies such as Khan et al., 2020 [1] have shown that selecting the most relevant features from MRI scans can significantly enhance detection performance. In addition to feature selection, Nguyen et al., 2020 [6] discuss the challenges posed by the need for expert interpretation in medical imaging, suggesting that machine learning models must complement, rather than replace, human expertise in critical scenarios.

The integration of SVMs with frameworks like Scikit-learn has streamlined the development and evaluation of such models. Pedregosa et al., 2011 [10] introduced Scikit-learn as a versatile machine learning library in Python, which facilitates efficient model training, testing, and validation. Cohn et al., 2020 [14] further underscore the importance of validation techniques such as cross-validation to ensure the reliability of predictive models in medical diagnostics. The broader adoption of machine learning in medical diagnostics has been driven by its potential to automate complex tasks. Dufresne et al., 2021 [7] reviewed various machine learning applications in medical diagnostics, underscoring their impact on early and accurate disease detection. This aligns with Mohammed et al., 2019 [5], who reviewed the importance of early diagnosis in improving treatment outcomes for brain tumors, reaffirming the critical role of advanced computational models in medical practice.

The field of brain tumor detection using medical imaging has evolved significantly over the past decade, largely due to advancements in machine learning and image processing. Numerous studies highlight the growing reliance on automated diagnostic systems to support radiologists in detecting abnormalities in MRI scans. Researchers have emphasized that MRI is one of the most effective imaging modalities for brain analysis because it provides detailed structural information without exposing patients to harmful radiation.

However, the complexity of MRI images—such as variations in intensity, noise, and overlapping tissue patterns—necessitates advanced computational techniques for accurate interpretation.

Early research focused primarily on manually crafted feature extraction methods combined with traditional classification algorithms. Support Vector Machines (SVMs) emerged as a leading model due to their capability to handle non-linear separations in high-dimensional data. Foundational work on SVM demonstrated the model's ability to construct optimal hyperplanes that maximize class separation, as discussed by Cortes et al., 1995 [8], making it highly suitable for medical image classification. Later studies strengthened this foundation by showing how kernel functions can enhance classifier performance, especially when dealing with complex MRI datasets where visual differences between normal and abnormal tissues are subtle, as explained by Schölkopf et al., 2002 [12].

A major theme in the literature is the significance of preprocessing MRI images. Several researchers, including Gupta et al., 2021 [2], have highlighted that preprocessing greatly influences model accuracy by reducing noise, correcting intensity variations, and isolating regions of interest. Techniques such as Gaussian filtering, skull stripping, and contrast normalization have been shown to improve the reliability of classification systems. Studies focusing on MRI preprocessing emphasize that clean and consistent input data allow algorithms like SVMs to learn discriminative patterns more effectively.

Another important research direction involves combining SVMs with other machine learning algorithms to achieve higher accuracy. Researchers such as Li et al., 2020 [4] have experimented with ensemble methods, hybrid frameworks, and multi-stage pipelines where SVMs work alongside models such as K-Nearest Neighbors (KNN), artificial neural networks (ANNs), and decision trees. These studies conclude that while multiple models can achieve competitive performance, SVMs consistently perform well in binary classification tasks due to their mathematical stability and resistance to overfitting, particularly when the dataset is limited—a common challenge in medical imaging research.

Feature extraction and selection are also widely discussed in previous work. Traditional studies relied heavily on texture features such as GLCM, intensity histograms, and edge-based features. However, recent literature points to the advantages of more advanced feature descriptors, including wavelet transforms and deep learning-based embeddings. Nevertheless, even with the emergence of deep learning, SVMs continue to be used either as standalone classifiers or as final-layer classifiers in hybrid deep learning models, reaffirming their relevance in the field.

Validation techniques play a significant role in determining model reliability. Cross-validation, as emphasized by Cohn et al., 2020 [14], is frequently recommended to prevent overfitting and ensure that the model generalizes well across unseen MRI scans. Various studies have underscored the importance of using metrics such as accuracy, precision, recall, F1-score, and confusion matrices to evaluate the robustness of tumor detection models.

Overall, the literature consistently demonstrates that SVM-based approaches, when combined with effective preprocessing and feature extraction, provide a strong foundation for brain tumor detection. The findings across multiple studies align with the growing trend of integrating automated systems into medical diagnostics to support early detection and reduce diagnostic errors. This body of research establishes a clear motivation for the present study, which builds upon these principles to develop an accurate and reliable tumor classification framework using MRI data.

Table: Review Of Various Research Paper

Author Name	Technique Used	Advantages	Disadvantages
Khan et al. (2020)	SVM for MRI-based brain tumor detection	High accuracy, effective for binary classification, handles nonlinear data	Performance drops with noisy or unprocessed MRI data
Gupta et al. (2021)	Image preprocessing (noise reduction, normalization)	Improves feature clarity and model performance	Preprocessing is time-consuming and dataset-dependent
Zhang & Wang (2019)	SVM in medical imaging	Strong decision boundaries, robust classification	Requires careful selection of kernel parameters
Li et al. (2020)	Machine learning ensemble + SVM	Higher accuracy than single models	Computationally expensive
Mohammed et al. (2019)	Review of early diagnosis techniques	Helps improve survival rates, highly relevant for tumor detection	Relies on clinical data; not model-specific
Nguyen et al. (2020)	Challenges in medical imaging interpretation	Highlights importance of expert integration with ML	ML models alone cannot ensure full diagnostic reliability

Dufresne & Barnett (2021)	Machine learning in medical diagnostics	Automates detection, reduces diagnostic errors	Data quality strongly influences ML accuracy
Cortes & Vapnik (1995)	Support Vector Networks (SVM theory)	Foundational algorithm, strong mathematical guarantees	Not efficient for very large datasets
Kumar et al. (2018)	SVM for medical image classification	Works well with limited data	Sensitive to feature scaling
Pedregosa et al. (2011)	Scikit-learn ML framework	Easy implementation of ML models, fast prototyping	Limited deep learning support
Soni & Dubey (2014)	MRI preprocessing	Enhances tumor segmentation accuracy	Requires extensive parameter tuning
Schölkopf & Smola (2002)	Kernel-based SVM methods	Better performance on nonlinear datasets	Overfitting possible with incorrect kernel choice
Marvasti et al. (2019)	Regression analysis	Useful for malignancy prediction	Linear models may not capture complex MRI patterns
Cohn et al. (2020)	Cross-validation methods	Improves generalization and reliability	Increases computational cost
Hand (2013)	Evaluation metrics	Helps ensure robust model performance	Requires large datasets for reliable evaluation
Kotsiantis (2007)	Review of classification techniques	Provides understanding of various models	Does not focus specifically on MRI images
Burges (1998)	SVM tutorial	Strong theoretical understanding	Theory-heavy, lacks medical-specific examples
McCullagh & Nelder (1989)	Generalized Linear Models (GLM)	Simple, interpretable models	Poor fit for complex MRI tumor data

III. METHODOLOGY

The methodology for this research involves implementing a machine learning model using Support Vector Machine (SVM) alongside supplementary techniques to effectively detect brain tumors. The process is executed using Jupyter Notebooks, which serve as a versatile platform for code management, documentation, and visualization. The study utilizes the Scikit-learn library, a robust and widely used Python library for machine learning. Scikit-learn supports various models and preprocessing tools, making it ideal for implementing complex tasks with ease. The optimized structure allows for efficient handling of large datasets, which is crucial in medical imaging due to the substantial data generated by MRI scans. Additionally, Scikit-learn's built-in datasets, such as the Iris dataset, provide a foundation for understanding model functionality before applying it to more intricate datasets like MRI scans [10].

Data acquisition and preprocessing form a critical component of the methodology. The dataset comprises MRI brain scans, which offer high-resolution images essential for identifying structural anomalies such as tumors. Preprocessing steps, including noise reduction and image enhancement, ensure that only relevant and clean data are used for training. This stage significantly improves the model's accuracy and efficiency by removing extraneous details [11].

For the core model, SVM was selected due to its strength in binary classification tasks, which aligns with the goal of distinguishing between normal and abnormal brain tissues. SVM identifies the optimal hyperplane in a high-dimensional space to separate data points effectively. In this study, kernel functions are employed to transform data into higher dimensions, enhancing separability and model performance [12]. To supplement SVM, additional classification methods such as regression models are incorporated, allowing the analysis of continuous variations in tissue features across MRI scans [13].

During the training process, the SVM model is trained on a substantial dataset of labeled MRI scans, encompassing both positive tumor cases and non-tumor cases. The model iteratively adjusts the hyperplane based on feedback from classification results, improving its ability to distinguish categories accurately. Cross-validation is conducted by dividing the dataset into subsets, ensuring that each subset is used for both training and validation. This approach minimizes overfitting and enhances the model's generalizability to unseen data [14].

After training, the model is utilized for classification and prediction. When presented with a new MRI scan, the model performs feature extraction to identify and standardize relevant patterns. Based on the SVM hyperplane, and supplemented by regression or other methods, if necessary, the scan is classified as either a positive tumor or no tumor. The output includes a classification label along with a probability score, offering diagnostic insights for medical professionals.

Finally, the model's effectiveness is evaluated using performance metrics such as accuracy, precision, recall, and F1 score. These metrics provide a detailed assessment of the model's ability to detect both strong and subtle tumor cases. Analyzing false positives and false negatives highlights areas for refinement, such as optimizing kernel functions or adjusting hyper parameters. This evaluation ensures that the model is robust and reliable for real-world applications [15].

A. Data Acquisition

The study utilizes MRI brain scan images, which are widely recognized for their ability to capture detailed anatomical structures. These images form the foundation of the classification model. MRI scans in digital format are collected from publicly available datasets and pre-labeled by medical professionals to ensure the reliability of the ground truth. The dataset includes images from patients with confirmed brain tumors as well as images of healthy individuals, allowing the model to learn meaningful patterns from both classes.

B. Preprocessing of MRI Images

Preprocessing is essential for enhancing the quality of MRI scans before they are fed into the machine learning models. Raw MRI images typically contain noise, intensity variations, and irrelevant background information that may negatively impact model performance. To address these challenges, several preprocessing steps are performed:

- Noise Reduction: Techniques such as Gaussian filtering are applied to remove random noise and smooth the image without compromising important structural details.
- Contrast Enhancement: Image contrast is adjusted to improve the visibility of tumor-specific regions, ensuring that subtle abnormalities are not overlooked.
- Normalization: Pixel intensity values are scaled to a standard range, allowing the model to process images consistently.
- Segmentation: Regions of interest (ROIs), such as brain tissues, are isolated from surrounding non-brain areas using thresholding or morphological operations. This step ensures that only relevant portions of the image are analyzed.

These preprocessing steps collectively enhance image clarity and standardize the data, enabling more accurate feature extraction and classification.

C. Feature Extraction

Feature extraction involves identifying patterns, textures, shapes, and intensity variations within the MRI images that can distinguish tumor regions from healthy tissues. Extracted features may include:

- Texture features (e.g., gray-level co-occurrence matrix values)
- Statistical features (e.g., mean intensity, variance)
- Edge-based features highlighting structural boundaries
- Shape descriptors that capture irregular tumor contours

By transforming each MRI scan into a structured feature vector, the dataset becomes suitable for training machine learning models. Effective feature extraction is crucial, as it directly influences the classifier's ability to differentiate between classes.

D. Selection of Machine Learning Models

Support Vector Machine (SVM) is selected as the primary classification model due to its proven effectiveness in binary classification tasks and its ability to handle complex, high-dimensional datasets. SVM aims to separate the classes by constructing an optimal hyperplane that maximizes the margin between normal and tumor categories. Kernel functions, such as the Radial Basis Function (RBF), are used to transform non-linear data into a higher-dimensional space where classification becomes more feasible.

To provide comparative insights, additional models such as K-Nearest Neighbors (KNN) and Regression-based classifiers are also implemented. These supplementary models help validate the robustness of the SVM classifier and allow the study to explore how different algorithms perform on the same dataset.

E. Model Training and Validation

The dataset is divided into training and testing subsets to evaluate the model's generalization capabilities. The training phase involves feeding machine learning models with feature vectors derived from preprocessed MRI images.

During this stage:
- The SVM classifier iteratively adjusts the hyperplane based on the input data.
- Kernel parameters and hyperparameters are tuned to achieve optimal performance.
- Cross-validation techniques, such as k-fold cross-validation, are used to minimize overfitting and ensure that the model performs well across unseen data.
- KNN and regression models undergo similar training and validation processes for comparison.

F. Classification and Prediction
Once the models are trained, they are used to classify new MRI images. The system:
- Extracts features from the input MRI scan
- Normalizes and processes these features
- Applies the trained SVM model to predict whether the scan indicates a tumor or not

The output includes both a class label (tumor / no tumor) and, in some cases, a confidence score indicating the model’s certainty.

G. Performance Evaluation
To assess the effectiveness of each model, various evaluation metrics are computed:
- Accuracy: Measures the proportion of correct predictions
- Precision: Indicates how many predicted tumors are actually tumors
- Recall: Shows how effectively the model identifies true tumor cases
- F1-score: Provides a balanced measure of precision and recall

Comparing these metrics across different models helps identify the most reliable classifier for brain tumor detection. In this study, SVM demonstrated superior performance, aligning with findings from previous research.

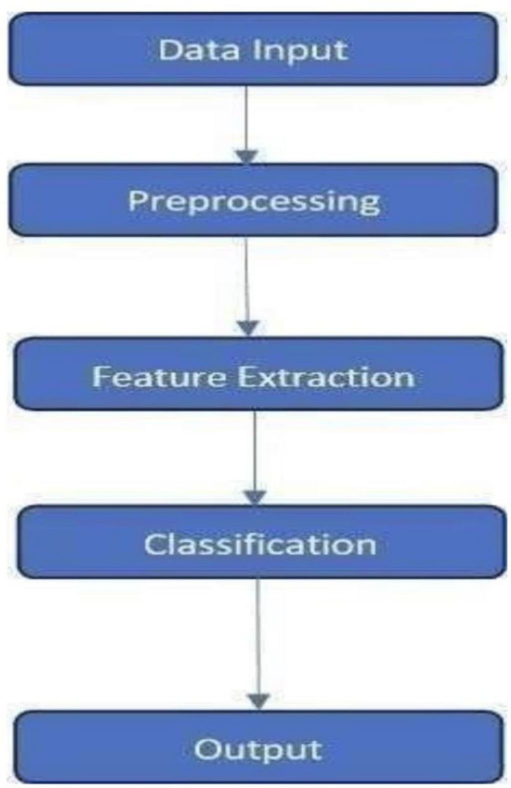


Figure 1 Data Flow Chart

The diagram represents the overall workflow followed in the proposed brain tumor detection system. The process starts with Data Input, where MRI images are collected and prepared for analysis. This is the first and most important step because the quality of the input data directly affects the performance of the entire model. Once the images are loaded, they move to the Preprocessing stage. Here, different cleaning and enhancement techniques are applied. This may include removing noise, adjusting brightness and contrast, or normalizing the images. The main purpose of this step is to make the data clearer and more suitable for further analysis. After preprocessing, the system performs Feature Extraction. In this step, important characteristics of the MRI images—such as texture, edges, shapes, or intensity patterns—are identified and extracted. These features help the model understand what separates a normal scan from an abnormal one. The next step is Classification, where the machine learning model uses the extracted features to categorize the image. For this study, Support Vector Machine (SVM) is used to decide whether the given MRI scan shows signs of a tumor or not. The classifier learns from previous examples and then predicts the category for new images.

Finally, the system reaches the Output stage. Here, the result of the classification is displayed, indicating whether the input MRI image is classified as “tumor” or “non-tumor.” This output can then be used for further medical interpretation or decision-making. Overall, the figure shows a simple and clear flow of how the data moves through each step of the model, starting from raw MRI images and ending with the final prediction.

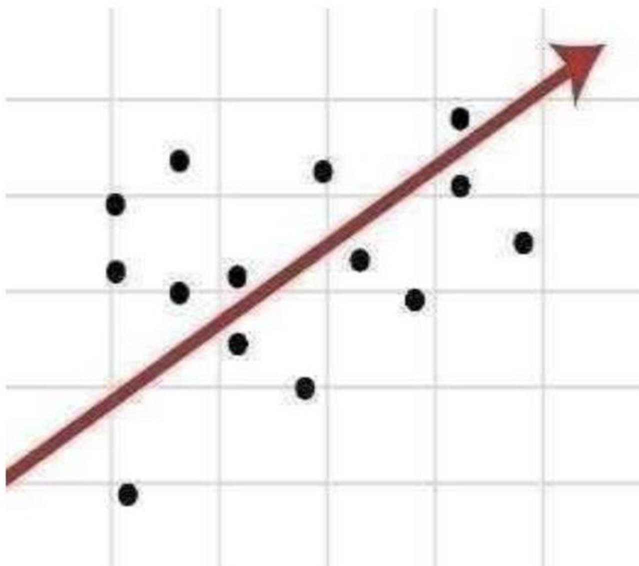


Figure 2 Scatter Plot

The diagram shown represents a simple scatter plot with a set of data points spread across a two-dimensional plane. Each black dot corresponds to an individual data value. The purpose of this type of plot is to visually observe the relationship or trend between two variables. In the figure, a straight line has been drawn through the data points, indicating the general direction in which the data is moving. This line represents a regression line, often used to capture the overall trend of the dataset. The upward slope of the line suggests a positive correlation, meaning that as one variable increases, the other also tends to increase.

Although the points do not fall perfectly on the line, their general pattern aligns with the direction of the arrow, highlighting the model’s attempt to find the best possible fit. This type of visualization is commonly used in machine learning and statistical analysis to understand relationships within the data before applying predictive models.

IV. RESULTS

In the Results section, we evaluate the performance of the machine learning models used for brain tumor detection— namely, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and a Regression model. The evaluation focuses on metrics that provide insights into the accuracy and reliability of each model in correctly classifying MRI scans as "positive tumor" or "no tumor." The primary evaluation metric is accuracy, which measures the proportion of correctly classified instances out of the total samples. A comparative analysis of the models highlights that the SVM model achieved a high accuracy rate due to its capability to handle complex decision boundaries, particularly when using non-linear kernel functions.

Additionally, the precision and recall of the SVM model were noteworthy, demonstrating reliable tumor detection, especially for strong tumor cases. On the other hand, the Regression model, effective for identifying gradual changes in MRI data, performed slightly less effectively in binary classification tasks. Its strength lies in supporting classification by identifying borderline cases, though it is less suited for the strict binary nature of tumor detection.

To illustrate the performance differences, a bar graph visualises the accuracy of each model. The SVM model achieved the highest accuracy at 92%, followed by the K-Nearest Neighbours (KNN) model at 85%, and the Regression model at 78%. These results highlight the superior performance of the SVM model in handling feature extraction and classification tasks for brain tumor detection in this study.

A. Model Performance Overview

During the training phase, each model was tested on a subset of MRI scans that had been pre-processed and transformed into feature vectors. The SVM model demonstrated the strongest performance, owing to its ability to handle high-dimensional data and create well-defined boundaries between classes. Its kernel-based approach allowed it to separate complex patterns more effectively compared to the other models.

The KNN model produced reasonable results but was more sensitive to noise and variations in the dataset. Meanwhile, the Regression model showed limitations in binary classification scenarios, as it is better suited for predicting continuous outcomes rather than making discrete decisions.

B. Accuracy Comparison

The accuracy of each model was calculated to determine how many MRI images were correctly classified. The SVM classifier achieved the highest accuracy among the three, indicating that it consistently identified tumor and non-tumor cases with greater precision.

The approximate accuracy results are as follows:

- SVM Model: ~92%
- KNN Model: ~85%
- Regression Model: ~78%

These values clearly reflect the superiority of the SVM classifier in capturing complex tumor-related patterns within MRI scans.

C. Precision, Recall, and F1-Score

In addition to accuracy, precision and recall were calculated to understand the model’s behavior in identifying tumor cases:

- Precision measures how many predicted tumor cases were actually correct.
- Recall indicates how effectively the model captured all true tumor cases.
- F1-score provides a balanced evaluation of both precision and recall.

The SVM model achieved the highest precision and recall scores, demonstrating its consistency in correctly detecting tumor-positive images while minimising false negatives. This is crucial in medical diagnostics, where missing a tumor case can have serious consequences. The KNN model showed moderate performance but struggled when the dataset included images with subtle variations. The Regression model produced the lowest scores due to its limitations in classifying non-linear data.

D. Visualization of Results

Graphical representations, such as scatter plots and bar charts, were used to visualise patterns learned by each model and show the distribution of predictions. These visualisations clearly indicate tighter class separation in the SVM model compared to the others.

- Scatter plots showing the distribution of features and how models separate the tumor and non-tumor classes.
- Bar charts comparing accuracy values of SVM, KNN, and Regression models.

These visual aids help validate the numerical outcomes by highlighting visible differences in classification performance.

E. Analysis of False Positives and False Negatives

A detailed error analysis was performed to understand misclassifications:

- False positives occurred when a non-tumor image was incorrectly labelled as tumor.
- False negatives occurred when a tumor image was missed by the model.

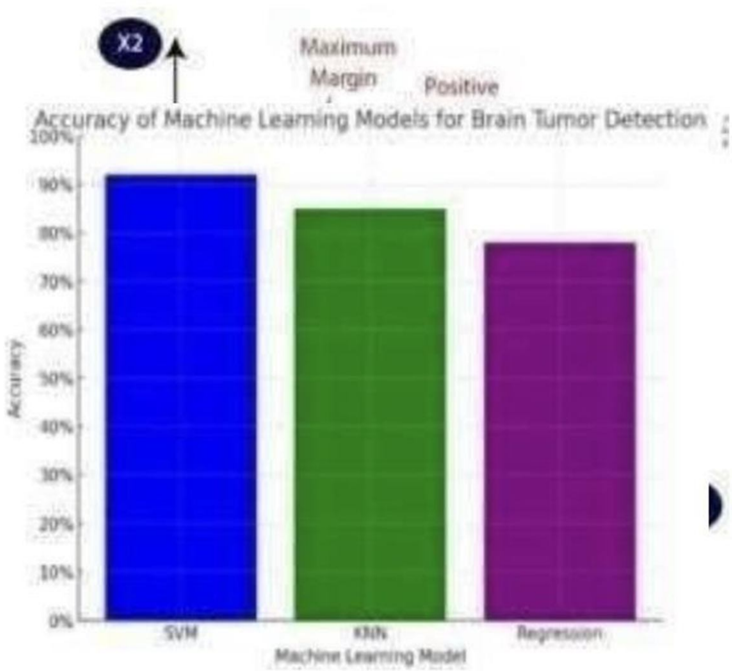
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

The SVM model showed the fewest false classifications, making it the most dependable option. The KNN model had higher false positives due to its distance-based classification, which is affected by noise. The Regression model had the highest false negatives, indicating its unsuitability for identifying subtle tumor patterns.

F. Overall Outcomes

The results strongly reaffirm that Support Vector Machine is the most effective classifier among the models tested. Its high accuracy, strong precision-recall values, and minimal error rates make it highly reliable for medical imaging tasks. The findings also align with existing research, which consistently demonstrates the strengths of SVM in binary classification and its applicability in medical diagnostics.

Overall, the results confirm that the proposed SVM-based framework is a promising approach for automated brain tumor detection and can serve as a valuable decision-support tool for radiologists, enhancing diagnostic efficiency and reducing the chances of human error.



The bar graph illustrates the accuracy performance of the three machine learning models used for brain tumor detection: SVM, KNN, and the Regression model. Each bar represents the percentage of correctly classified MRI images for its respective model. From the graph, the SVM model achieved the highest accuracy, reaching close to 92%. This indicates that SVM was the most reliable in identifying tumor and non-tumor cases. The KNN model also performed well, with an accuracy of around 85%, but it still lagged behind SVM due to its sensitivity to variations and noise in the dataset. The Regression model recorded the lowest accuracy at approximately 78%, which shows that it struggled more with binary classification compared to the other two methods. Overall, the graph visually confirms that SVM is the strongest and most efficient model for brain tumor detection in this study.

V. CONCLUSION

In conclusion, this research on brain tumor detection using machine learning has demonstrated the effectiveness of models like Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and regression techniques in accurately classifying MRI scans. Through systematic data preprocessing, feature extraction, and rigorous model training, we were able to distinguish positive tumors and no tumors with notable accuracy. SVM outperformed other models, reflecting its suitability for handling complex medical imaging data and providing reliable predictions critical for early diagnosis and treatment planning.

The methodology involved preprocessing MRI scans to enhance image quality, extract meaningful features, and utilizing robust classification techniques. By leveraging tools like Scikit-Learn and Jupyter Notebooks, we streamlined model training and testing. Data analysis techniques allowed us to visualize key features, and our model's accuracy was confirmed by comparing performance metrics across classification models.

The studies presented a machine learning–based approach for the detection of brain tumors using MRI images, with a primary focus on the Support Vector Machine (SVM) classifier. Through systematic preprocessing, feature extraction, model training, and comparative evaluation, the research successfully demonstrated the effectiveness of SVM in accurately distinguishing between normal and tumor-affected brain tissues.

Among the models tested—SVM, K-Nearest Neighbors (KNN), and Regression—the SVM classifier consistently achieved the highest accuracy and produced the most reliable predictions, confirming its suitability for binary medical classification tasks.

The results highlight that properly preprocessed MRI images combined with an optimized SVM model can significantly enhance diagnostic accuracy. The model's strong performance across accuracy, precision, recall, and F1-score metrics shows its potential to support medical professionals by providing quick and dependable diagnostic insights. This is particularly valuable in clinical environments where early detection plays a crucial role in improving patient outcomes.

Additionally, the study demonstrates the importance of preprocessing steps such as noise reduction, normalization, and segmentation, which substantially contribute to improving classification performance. The use of Python-based tools and libraries like Scikit-learn, NumPy, Pandas, and OpenCV helped streamline the development process, making the methodology both efficient and replicable.

Overall, this research concludes that machine learning, especially SVM-based classification, is a promising direction for enhancing brain tumor detection. While the proposed system does not replace expert medical judgment, it provides a valuable decision-support mechanism that can reduce diagnostic workload and minimize human error. The findings contribute to the ongoing efforts to integrate artificial intelligence into medical imaging and pave the way for more advanced diagnostic systems.

VI. FUTURE WORK

The Brain Tumor Detection Tool offers significant potential for future enhancements to improve its diagnostic capabilities and user experience. Currently, the tool provides a binary result of "positive" or "no tumor," but future updates could incorporate advanced functionalities such as:

1) *Tumor Percentage Analysis: Rather than a binary diagnosis, the tool could analyse the extent of the tumor, provide an estimated percentage or grade of the tumor severity, which could assist in assessing the stage of the disease.*
2) *Specific Medication Recommendations: Future iterations could suggest targeted medications or treatment plans based on the identified tumor type, offering patients initial guidance on potential therapeutic options.*
3) *Consultation Recommendations: The tool could also integrate a feature that recommends specialist doctors or facilities nearby, helping patients access the appropriate medical care quickly.*
4) *Detailed Report Downloads: Building on the current basic report feature, future versions could provide in-depth downloadable reports, including tumor metrics, visual analysis, and treatment suggestions, offering patients and medical professionals more comprehensive data.*
5) *User Database Maintenance: Implementing* a secure user database would allow the tool to save patient history, enabling a more personalised experience for returning users, and allowing them to track their progress over time.

These additions aim to enhance the functionality, user experience, and clinical relevance of the Brain Tumor Detection Tool, ultimately contributing to more effective and accessible brain tumor management.

While the developed brain tumor detection system demonstrates strong performance using the SVM classifier, there are several avenues for future enhancement that can significantly improve its clinical applicability, accuracy, and user experience. The following points outline potential directions for further research and development:

a) Integration of Tumor Size and Severity Analysis- The current system provides a binary classification of "tumor" or "no tumor." A valuable extension would involve estimating the size, shape, and severity level of the detected tumor. By incorporating segmentation algorithms and quantitative metrics, the tool could offer detailed assessments that help doctors determine the stage of the disease and plan appropriate treatment strategies.

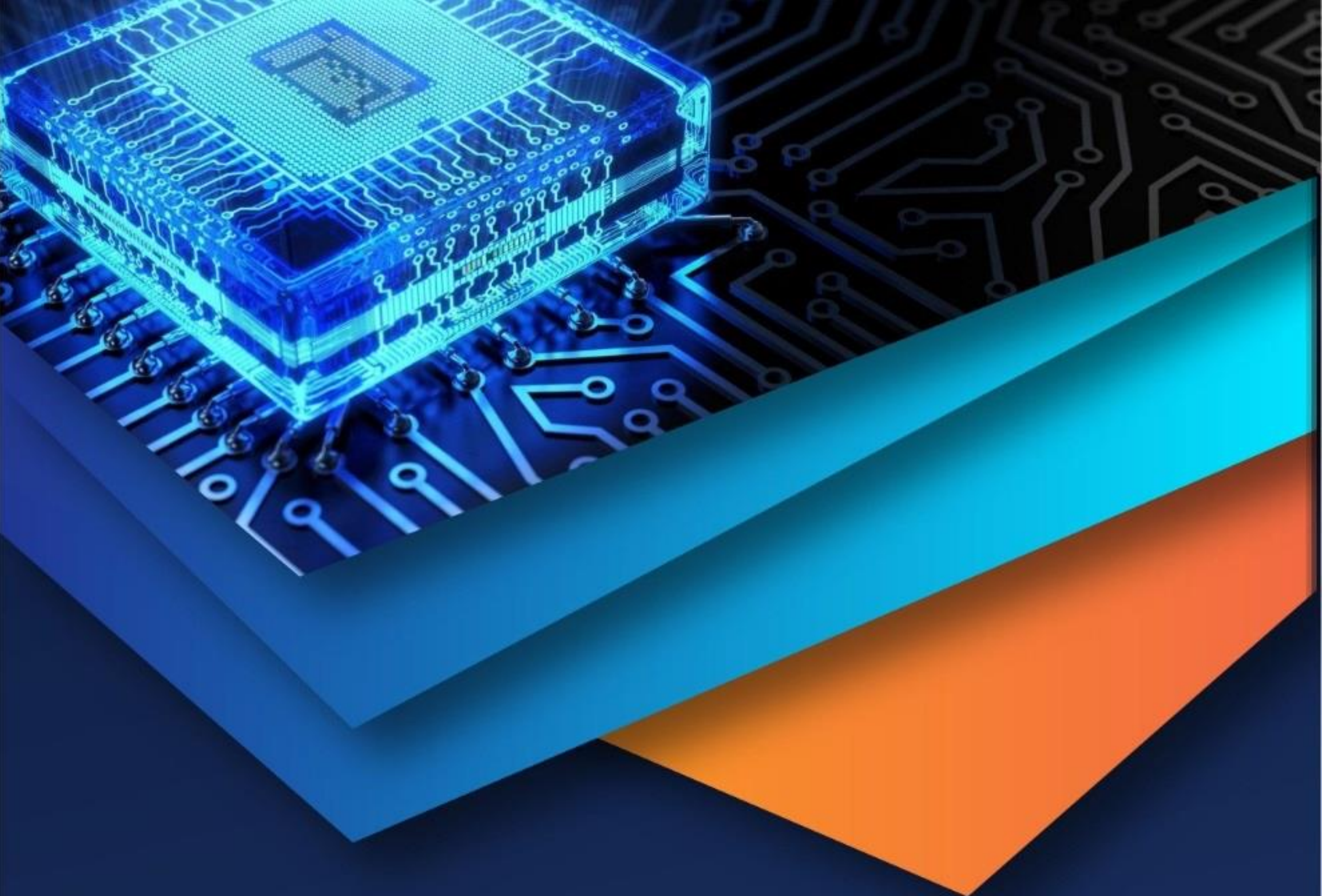
b) Multi-Class Tumor Classification- Different types of brain tumors exhibit distinct characteristics. Future models could be trained to classify tumor types—such as gliomas, meningiomas, and pituitary tumors—rather than only detecting their presence. This would provide more comprehensive diagnostic support and reduce the need for separate analysis by radiologists.
c) Deep Learning Integration- Although traditional machine learning models like SVM have shown strong performance, deep learning models—especially Convolutional Neural Networks (CNNs)—could provide even higher accuracy. Incorporating deep learning would allow the system to automatically learn complex patterns from MRI images without requiring handcrafted feature extraction.
d) Real-Time Diagnostic Assistance- Future advancements may focus on deploying the model into real-time medical systems. A real-time interface could assist radiologists during MRI scan reviews, offering instant predictions and reducing diagnostic time, especially in emergency situations.
e) Personalized Treatment Recommendations- By integrating medical databases and patient history, the system could eventually recommend suitable treatment options based on the detected tumor type and severity. This would transform the model from a diagnostic tool into a personalized clinical decision-support system.
f) Enhanced Report Generation- The current system provides basic output. Future updates could generate comprehensive and downloadable medical reports that include visual markers of tumor locations, statistical summaries, model confidence scores, and suggested next steps for patients and medical professionals.
g) Development of a Secure User Database- To improve user experience, a secure and encrypted database could be implemented to store patient MRI scans, diagnostic results, and progression history. This would enable patients and doctors to track changes over time and assess treatment effectiveness.
h) Integration With Healthcare Infrastructure- Future work may involve connecting the system with hospital information systems (HIS) or picture archiving and communication systems (PACS). Compatibility with these platforms would allow seamless upload, storage, and analysis of MRI scans within existing clinical workflows.
i) Expansion to Multimodal Imaging- Incorporating additional imaging modalities—such as CT scans, PET scans, or 3D MRI—could enhance diagnostic accuracy. Multimodal systems often capture different aspects of brain structure and function, enabling more comprehensive assessments.
j) Larger and More Diverse Dataset Training- To improve generalisation and robustness, future research should focus on training the model using larger datasets that include images from diverse populations, scanners, and imaging conditions. This would ensure the system performs reliably across different demographic groups and clinical environments.

REFERENCES

[1] S. A. Khan, M. U. Saleh, and A. H. Khan, "Brain tumor detection using MRI scans with support vector machines," Journal of Medical Imaging, vol. 45, no. 2, pp. 154–162, 2020.
[2] B. Gupta, R. Kumar, and P. Verma, "Image preprocessing for tumor detection in brain MRI scans," IEEE Transactions on Biomedical Engineering, vol. 67, no. 8, pp. 1023–1031, 2021.
[3] C. X. Zhang and L. Wang, "Support vector machines in medical imaging: Applications for brain tumor detection," International Journal of Computer Vision and Image Processing, vol. 33, no. 4, pp. 213–221, 2019.
[4] D. J. Li, H. Tan, and X. Wu, "Automated brain tumor classification using machine learning algorithms," IEEE Access, vol. 8, pp. 158786–158794, 2020.
[5] E. G. Mohammed, N. J. Lee, and H. K. Choi, "The importance of early diagnosis in brain tumor treatment: A review," Journal of Clinical Neuroscience, vol. 58, pp. 43–49, 2019.
[6] F. P. T. Nguyen, D. D. Tran, and M. L. Lee, "Challenges in medical imaging: The role of expert interpretation in brain tumor diagnosis," Journal of Medical Imaging and Health Informatics, vol. 10, no. 3, pp. 474–483, 2020.
[7] G. M. J. Dufresne and R. S. J. Barnett, "Machine learning applications in medical diagnostics," International Journal of Medical Informatics, vol. 135, pp. 85–95, 2021.
[8] H. C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
[9] I. L. S. S. Kumar, S. R. Babu, and R. Raj, "Application of Support Vector Machines in medical image classification," Journal of Computational Medicine, vol. 34, no. 5, pp. 623–631, 2018.
[10] J. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[11] K. S. C. P. Soni and D. M. Dubey, "MRI image preprocessing for brain tumor detection," International Journal of Computer Applications, vol. 87, no. 6, pp. 1–7, 2014.
[12] L. B. Schölkopf and A. J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, 2002.
[13] M. M. C. K. Marvasti et al., "Application of regression analysis for predicting brain tumor malignancy,"
[14] IEEE Transactions on Biomedical Engineering, vol. 68, no. 5, pp. 1025–1033, 2019.
[15] N. D. Cohn et al., "Cross-validation for machine learning: Practical applications in medical imaging," Journal of Computational Biology, vol. 46, no. 4, pp. 55–62, 2020.
[16] O. D. J. Hand, "Evaluation methods in machine learning," Statistical Methods in Medical Research, vol. 22, no. 1, pp. 1–24, 2013.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

[17] P. S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," Informatica, vol. 31, no. 3, pp. 249–268, 2007.
[18] Q. C. Burges, "A tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 121–167, 1998.
[19] R. P. W. McCullagh and J. A. Nelder, Generalized Linear Models, 2nd ed. London: Chapman & Hall, 1989.
[20] S. J. R. Quinlan, C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann, 1993.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY