

Building a brain cancer treatment hospital recommendation system in Vietnam

Truong Ho-Viet Phan¹, Anh Nhat Lam¹, Bao The Phung²

¹Faculty of Information Technology, School of Technology, Van Lang University, Ho Chi Minh City, Vietnam

²Faculty of Information Technology, Ho Chi Minh City University of Industry and Trade, Vietnam

Article Info

Article history:

Received Aug 8, 2025

Revised Dec 27, 2025

Accepted Apr 16, 2026

Keywords:

Brain cancer

Hospital recommendation

EfficientNet-B3

Neural network

Collaboration filtering

ABSTRACT

The purpose of this paper is to suggest an integrated system that combines the classification of medical images and a recommendation model for the choice of the best hospitals for cancer treatment according to the analysis of the brain MRI. For the classification of brain MRI images for tumor detection, we used EfficientNet-B3 as the base model. To evaluate its performance, we contrasted it with two other popular deep learning architectures, namely Vision Transformer (ViT) and ResNet101. Having identified malignancies by image classification, we used the outputs as inputs for a collaborative filtering based recommendation system. This system was constructed utilizing neural network embeddings and hidden vectors to learn the correlations between tumor types and hospital treatment characteristics. Personalized hospital suggestions were generated by calculating the similarity between the embeddings using the dot product. In addition, comparison evaluations with a traditional K-Nearest Neighbors (K-NN) method are performed to evaluate the performance gaps. The classification results revealed that EfficientNet-B3 has better accuracy when compared with ViT and ResNet101 and it is a good choice for medical picture analysis. The neural collaborative filtering model shows superior accuracy and suitability of the hospital selection during the recommendation phase compared to K-NN. The results of this study demonstrate that the combination of deep learning and intelligent recommendation systems may successfully provide excellent diagnosis in healthcare. The proposed framework may improve decision making and may provide individualized recommendations of hospitals for cancer treatment. A case study is offered in Vietnam to recommend hospitals.

Copyright © 2026 Institute of Advanced Engineering and Science.

All rights reserved.

Corresponding Author:

Bao The Phung

Faculty of Information Technology, Ho Chi Minh City University of Industry and Trade, Vietnam

140 Le Trong Tan street, Tay Thanh Ward, Ho Chi Minh City, Vietnam

Email: baopt@huit.edu.vn

1. INTRODUCTION

Recommendation system is widely used in many sectors. The most common one is the movie suggestion system in online movie viewing platforms [1, 2]. Today, the use of recommendation systems is more common than ever but remains rare in the medical profession [3–7]. Therefore, the combination of recommendation systems with computer vision for medical picture classification creates a new research avenue for recommendation systems. Generally, recommendation systems are categorized into three basic forms, which are collaborative filtering (CF), content based filtering (CBF) and hybrid [8]. In CF, the system is separated into two recommendation directions, one is user-based, the system evaluates the similarity between users based on rating history, and recommends similar users' favorite goods in the same category [9, 10]. Using item-based,

the algorithm finds similar things to the items the consumers rated highly [11]. Permana addresses the topic of identifying the best collaborative filtering approach for restaurant recommendation systems [12]. Specifically, it compares User-Based Collaborative Filtering (UBCF) and Item-Based Collaborative Filtering (IBCF) with Cosine Similarity and Pearson Correlation to show which combination can provide the most accurate and dependable recommendations to users. The accuracy and ranking consistency were poorer in Item-Based methods (IBCF). Guan [13] proposed to combine collaborative filtering and sentiment analysis to construct an effective location-based map recommender system. It considers the limitations of merely considering the numerical ratings, by taking into account text review analysis and location-aware techniques, thus allowing for more personalized and context-aware recommendations. The author also surveys and describes models such as Singular Value Decomposition (SVD), Alternating Least Squares (ALS), K-Nearest Neighbors (KNN), and sentiment scoring approaches like VADER, and how they might be used together. Poudel addresses the issue of knowing how the structure of rating data affects the effectiveness of collaborative filtering (CF) algorithms [13]. Prior study has employed six or more data features to explain the differences in CF performance. Poudel aims to find out whether a smaller number of critical characteristics may perform as well or better. Specifically, we evaluate if the number of ratings per user (Information per User, IpU) and number of ratings per item (Information per Item, IpI) are adequate to explain most of the diversity of CF performance among the different methods and datasets.

The recommendation system for CBF is based on the user's profile or the features of items similar to those already selected by the user [14]. Tian focuses on the task of generating personalized movie recommendations with content-based filtering (CBF) algorithms [15]. It builds a content-based recommendation system using movie metadata and user profile features to generate relevant recommendations, addressing common limitations of collaborative filtering, such as the cold start problem, inadequate user ratings, and over-reliance on interaction data. They propose to enhance CBF with keyword pre-processing and weighted rating calculations to achieve higher accuracy and diversity in the recommendations. The overchoice problem is a cognitive problem that happens when people have too many equal options to choose from, a situation that is common in places such as cafes. Kosim et al treat the subject [16]. This paper attempts to address this issue at Muftada Kopi, a café in Indonesia. We design a beverage recommendation system that is non-personalized by using content-based filtering (CBF). The aim is to facilitate the user's choice of drinks by linking his declared preferences to the features of the drinks, improving the efficiency of decision making and minimizing the waiting time for service. Anh and Truong tackle the issue of aiding IT students in securing appropriate employment prior to graduation by aligning their academic competencies with employer expectations [17]. Many students are unable to fulfill job criteria due to insufficient experience or inadequate skills. In this study, we propose a content-based recommendation system that matches students' academic performance (grades in major courses) to job postings from IT companies to find the best matching jobs for the student capabilities. The tool provides a job discovery and allows the students to identify the skills gap early, thus increasing the students' readiness for the labor market ahead of graduation.

The recommendation system for CBF is based on the user's profile or the features of items similar to those already selected by the user [14]. Tian focuses on the task of generating personalized movie recommendations with content-based filtering (CBF) algorithms [15]. It builds a content-based recommendation system that uses movie metadata and user profile features to produce relevant recommendations. It solves common problems of collaborative filtering such as the cold start problem, not having enough user ratings and over-dependence on interaction data. They propose to enhance CBF with keyword pre-processing and weighted rating calculations to achieve higher accuracy and diversity in the recommendations. The overchoice problem is a cognitive problem that happens when people have too many equal options to choose from, a situation that is common in places such as cafes. Kosim et al treat the subject [16]. This paper attempts to address this issue at Muftada Kopi, a café in Indonesia. We design a beverage recommendation system that is non-personalized by using content-based filtering (CBF). The aim is to facilitate the user's choice of drinks by linking his declared preferences to the features of the drinks, improving the efficiency of decision making and minimizing the waiting time for service. Anh and Truong tackle the issue of aiding IT students in securing appropriate employment prior to graduation by aligning their academic competencies with employer expectations [17]. Many students are unable to fulfill job criteria due to insufficient experience or inadequate skills. In this study, we propose a content-based recommendation system that matches students' academic performance (grades in major courses) to job postings from IT companies to find the best matching jobs for the student capabilities. The tool provides a job discovery and allows the students to identify the skills gap early, thus increasing the students' readiness

for the labor market ahead of graduation.

However, there are still quite a few studies on combining computer vision and recommendation system to classify medical images and give treatment hospital suggestions. Therefore, in this paper, we propose a solution to build a hospital recommendation system through classifying brain MRI images. This study uses a family of convolutional neural networks called EfficientNet-B3 to recognize and classify brain MRI images. Then use that classification data as input to the neural network based collaboration filtering to suggest the appropriate hospital for treatment.

2. METHOD

In this section, we describe in detail the steps of building the system as shown in Fig. 1, our system consists of three main parts: MRI image dataset, hospital information dataset, MRI image classification process and finally the hospital recommendation filtering. In the brain MRI images module, we present how to collect and process MRI images. In the medical image classification module, we present EfficientNet-B3, a family of Convolutional Neural Network, used to classify brain MRI images according to 4 different labels, respectively glioma_tumor, meningioma_tumor, no_tumor and pituitary_tumor. In hospital dataset, we describe the storing structure for recommendation. In the collaboration filtering, we use an algorithm of collaborative filtering based on deep learning method to make appropriate hospital list based on the output of the MRI image classification module.

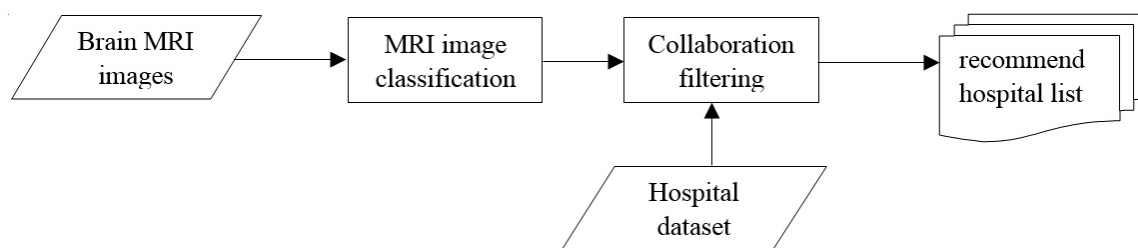


Figure 1. The overall diagram of our proposed method

2.1. Brain MRI images

For research purposes, our system uses brain MRI image dataset obtained from kaggle website. This dataset contains 7023 images of human brain MRI images which are classified into 4 classes: glioma - meningioma - no tumor and pituitary. After data processing and data augmentation, we get 10,000 images for each label, which is a total of 40,394 images for the entire training data for the MRI image classification module. Of which, 80% of the images are used for training, 20% are for validation and about 100 images per label are reserved for testing. Table 1 summarizes the number images of four tumor classes.

Table 1. Summarize the number of image for 4 tags of brain MRI images from kaggle.com

Classes	Train	Test
glioma_tumor	8,000	2,100
meningioma_tumor	8,000	2,115
no_tumor	8,000	2,105
pituitary_tumor	8,000	2,078

It is important to note that this dataset is publicly available and not specifically collected from Vietnamese hospitals. Due to limitations in accessing local clinical MRI data, we used this dataset to develop and evaluate the classification model. However, the proposed framework is designed to be adaptable and can be fine-tuned with local Vietnamese clinical data in future work.

2.2. Hospital dataset

Regarding hospital recommendation, we collected a total of 61 hospitals in Ho Chi Minh City that can meet the treatment and examination needs for brain cancer[27]. We used the 83 criteria scores in 2023

published on reputable Vietnamese online newspapers to make a rating, and searched for information: address, phone number, website on the official websites of each hospital as shown in Table 2.

Table 2. Hospital information storing structure for recommendation

ID	Hospital name	Address	Phone	Website	Rating
1	Hung Vuong hospital	179 Su Van Hanh, Ward 5, District 5	028 3924 2490	https://www.bvhungvuong.vn/	4.73
2	115 People hospital	527 Su Van Hanh, Ward 12, District 10	028 3864 2533	https://benhvien115.com.vn/	4.63
...	...	...	...	...	...
61	Saigon Tam Tri hospital	171 Truong Chinh street, Tan Son Nhi Ward, Tan Phu district	028 6293 4555	https://tmt.vn/	3.18

2.3. MRI image classification

In this section, we present the MRI image classification process as shown in Fig. 2. This process has three stages: i) input brain MRI images into the system and preprocess the data to ensure a consistent format, ii) EfficientNet-B3 model extracts image features through convolutional blocks and MBConv blocks, iii) the model output is passed through a fully connected and softmax layer to predict tumor labels and map them to tumor_id.

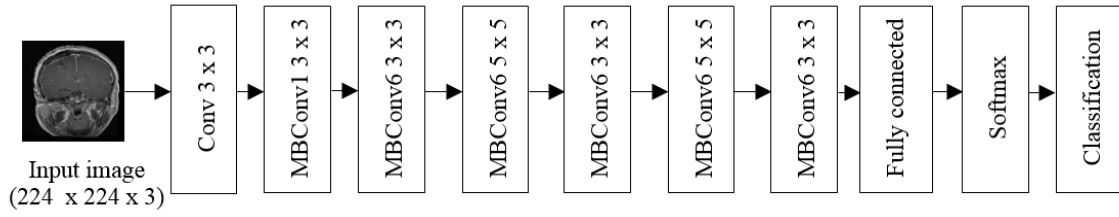


Figure 2. Classification process of EfficientNet-B3 with given brain tumor image

The EfficientNet-B3 is a CNN (Convolutional Neural Network) model of the EfficientNet family designed by Google AI [21]. It is known for high accuracy with less parameters and computations compared to previous models such as ResNet or VGG [22]. EfficientNet models scale up CNNs in a balanced way using a method called compound scaling, which uniformly scales depth, width, and resolution using a fixed set of coefficients. The core idea of EfficientNet is compound model scaling, which optimizes how the network should be made deeper, wider, and use higher-resolution input images with all in a balanced manner, instead of scaling each dimension arbitrarily. EfficientNet is built using MBConv (Mobile Inverted Bottleneck Convolution) blocks, enhanced with Squeeze-and-Excitation (SE) modules for better channel-wise attention. MBConv block was represented as [22]:

Expansion (1 x 1 convolution):

$$X_{exp} = ReLU6(X * W_{exp}), W_{exp} \in R^{1 \times 1 \times C \times tC} \quad (1)$$

Where, X is input feature map of shape $H \times W \times C$, W_{exp} is weights of the 1×1 convolution for expansion, $ReLU6$ is an activation function, X_{exp} is output with increased channels.

Depthwise Conv was computed as [22]:

$$X_{dw} = ReLU6(X_{exp} * W_{dw}), W_{dw} \in R^{k \times k \times tC \times 1} \quad (2)$$

Where, X_{exp} is output from expansion layer, W_{dw} is depthwise convolution filter of shape $k \times k \times tC \times 1$. K is kernel size (3 or 5), X_{dw} is output of depthwise conv, same spatial shape but still tC channels.

Squeeze-and-Excitation (SE) module includes Squeeze, excitation, and scale [28]. Squeeze is global average pooling to get channel descriptor:

$$z_c = \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W X_{dw}[i, j, c] \quad (3)$$

Where, z_c is pooled (squeezed) value for channel c , $X_{dw}[i, j, c]$ is value at spatial position (i, j) and channel c . H, W are height and width of feature map.

Excitation are two fully connected layers with non-linearity:

$$s = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot z)) \quad (4)$$

Where, z is vector of pooled values, shape $tC \times 1$, W_1 is the first fully connected layer, shape $\frac{tC}{r} \times tC$, r is the reduction ratio, W_2 is the second fully connected layer, shape $tC \times \frac{tC}{r}$.

Scale is the scaling the feature:

$$X_{se} = X_{dw} \cdot s \quad (5)$$

Where, X_{dw} is output from depthwise conv, s is channel-wise attention values, X_{se} is attention-weighted feature map.

The ability of EfficientNet-B3 in extracting the features is good and it leads to a better performance in image processing works. In this paper, EfficientNet-B3 is used to classify the brain MRI images into four classes such as glioma tumor, meningioma tumor, no tumor and pituitary tumor. 80% of the images are taken for training the images and they are preprocessed by downscaling them to 224×224 , normalization and applying data augmentation technique to improve the dataset diversity and reduce the overfitting. After data preparation, the images are passed to EfficientNet-B3, which extracts the image features using a series of MBConv blocks. Each MBConv block comprises an expansion convolution layer followed by a depthwise convolution layer, a squeeze-and-excitation operation and finally a projection convolution; skip connections are used when the dimensions of the input and output are matched which helps to solve the vanishing gradient problem [23]. The high dimensional feature maps produced from these blocks are processed by fully connected layers and then softmax algorithms are applied to obtain the probability distribution over the tumor categories and finally the predicted label is obtained.

2.4. Hospital Recommendation Based on Content-Based Filtering

In this research, a content-based filtering (CBF) approach is applied to choose appropriate hospitals based on the projected tumor type from MRI images [26]. Unlike collaborative filtering, which leverages the data of the user-item interaction, content-based filtering utilizes the intrinsic properties of the objects and the input features. This is especially true in medical settings, when explicit user feedback is sometimes not available due to privacy concerns and data limits. The output of the brain tumor classification algorithm to predict the label of a brain MRI picture is associated with a tumor_id. This tumor_id is then mapped to a feature representation via an embedding layer, which captures the semantic aspects of each type of tumor. Meanwhile, for the hospital data from the files hospital_info.csv and hospital_rating.csv, we employ another embedding layer to project each hospital_id to a feature vector which encodes the hospital's medical skills, specializations and quality metrics.

In our framework, each clinical case is represented by a feature vector derived from the predicted tumor type. Let $\mathbf{x} \in \mathbb{R}^d$ denote the feature representation of the input condition (i.e., tumor type). Each hospital i is also represented by a feature vector $\mathbf{y}_i \in \mathbb{R}^d$, which encodes relevant attributes such as medical specialization, treatment capability, available equipment, and hospital quality scores.

The relevance between the input condition and hospital i is computed using a similarity function:

$$s(i) = \text{sim}(\mathbf{x}, \mathbf{y}_i) \quad (6)$$

In this study, we employ cosine similarity due to its effectiveness in measuring the alignment between feature vectors:

$$s(i) = \frac{\mathbf{x} \cdot \mathbf{y}_i}{\|\mathbf{x}\| \|\mathbf{y}_i\|} \quad (7)$$

Based on the computed similarity scores, hospitals are ranked in descending order, and the top-k hospitals with the highest scores are selected as recommendations:

$$R_k = \text{Top}K\{s(i)|i \in H\} \quad (8)$$

where H denotes the set of all candidate hospitals.

This formulation allows the system to match clinical features to the capabilities of the hospitals quickly and without the need for historical interaction data. Hence, it is more robust in cold-start conditions and is more ideal for actual medical applications where individual feedback data is limited or absent.

The embedding vectors are learned in a supervised manner with the hospital rating data as the ground truth to provide significant representations for tumor types and facilities. For each tumor-hospital pair, the hospital dataset provides a target appropriateness score. This number indicates the overall quality and treatment skills of the hospital. Let y_i be the true rating of hospital i . Let $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{y}_i \in \mathbb{R}^d$ be the embedding vectors of the tumor type and the hospital i , respectively. The model derives a predicted appropriateness score as the dot product of these embeddings:

$$z_i = \mathbf{x} \cdot \mathbf{y}_i \quad (9)$$

The resulting score is then normalized to the range $[0,1]$ using the sigmoid function and scaled to a 5-point rating system:

$$\hat{y}_i = \sigma(z_i) \times 5 \quad (10)$$

where $\sigma(\cdot)$ denotes the sigmoid activation function. This formulation allows the model to interpret the dot product as a predicted rating that reflects the compatibility between a tumor type and a hospital.

To train the model, we minimize the Mean Squared Error (MSE) loss between the predicted scores and the ground truth ratings:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (11)$$

where N is the number of training samples. By optimizing this objective, the model learns embedding vectors such that their similarity aligns with real-world hospital ratings. This ensures that the learned hidden representations are not arbitrary, but are grounded in meaningful supervision derived from domain-specific data. Furthermore, the proposed approach can be naturally extended by incorporating additional clinical features or contextual information into the feature representation $\mathbf{x}$, thereby improving the flexibility and scalability of the recommendation system.

3. RESULT AND DISCUSSION

3.1. Settings

The training trials of EfficientNet-B3 model were performed on a personal computer equipped with an AMD R7-4800H CPU (2.9 GHz to 4.2 GHz, 8 cores, 16 threads), a GTX 3060 6 GB graphic card and 16 GB RAM. The model was trained with a maximum of 100 epochs with early termination mechanism to avoid overfitting and save computational power. We applied a neural network based collaborative filtering strategy with two hidden layers, with 128 neurons in the first layer and 32 neurons in the second layer. The output layer outputs the expected value for the tumor_id for the hospital_id. The neural network was trained over 5 epochs. We used k-fold cross-validation on the MRI classification dataset to get a more reliable estimate of model generalization. The dataset was partitioned into $k = 5$ folds, with $k-1$ folds used for training and 1 fold used for testing at each iteration. We report the average accuracy and standard deviation over all folds. This approach reduces the dependency on a single train/validation/test split and allows for a more robust evaluation of model stability.

This study utilizes a collection of publicly accessible brain MRI scans that have been completely anonymised, devoid of any personally identifiable information. All tests are performed offline and do not entail the gathering or storage of real-time patient data. Consequently, the study adheres to overarching data protection principles, encompassing the General Data Protection Regulation (GDPR) and pertinent Vietnamese data privacy laws.

3.2. Metrics

MAE is a commonly used metric to measure the accuracy of a model, especially in regression problems. MAE is the average of the absolute differences between predicted values and actual values. It tells, on average, how much the predictions are off from the actual results [27]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (12)$$

where y_i is the actual rating of hospital i , $\hat{y}_i$ is the model's predicted rating, and n is the total number of predictions.

Root Mean Square Error (RMSE) is another popular metric for evaluating the performance of regression models. RMSE is the square root of the average of the squared differences between predicted and actual values. Like MAE, it measures the differences between predicted and actual values but it gives more weight to large errors. RMSE is computed as [27]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (13)$$

Where, n is the number of tested brain images, y_i is the actual value of the hospital, $\hat{y}_i$ is the predicted value of the hospital.

3.3. Comparison of MRI brain image classification with EfficientNet-B3, ViT and ResNet101

We used three different models for the brain MRI image categorization experiment: Vision Transformer (ViT), ResNet101 and EfficientNet-B3. This choice is driven by the fact that there are three different approaches in Deep Learning, ViT uses a self-attention mechanism [30], ResNet101 makes use of skip connections [29] and EfficientNet-B3 adopts a compound scaling to optimize the network's depth, width and resolution [22]. It can be shown that the training speed of the ViT is the fastest during the training process. However, the final results show that the accuracy of the ViT is the lowest, with 2 of the 4 labels having less than 50% accuracy, and it is easy to confuse the features of different labels. ResNet101 took longer to train, but improved the accuracy of four labels significantly, and one label was still very different from other labels. The improved efficiency of EfficientNet-B3 showed better results in both temporal and accuracy issues, with an average accuracy of 85.38%. ResNet101 showed a higher training speed, but a lower average accuracy of 83.34% over four labels.

Table 3. The MRI brain image classification performance comparison of 3 models EfficientNet-B3, ViT and ResNet101

No	Models	Time (Hours)	Train loss	Accuracy
1	ViT	3.47	0.0296	0.6221
2	ResNet101	3.70	0.0301	0.8134
3	EfficientNet-V2	1.65	0.0084	0.9953
4	ConvNext	3.14	0.0043	0.9967
5	Swin	4.27	0.0073	0.9983
6	EfficientNet-B3	3.47	<u>0.0028</u>	<u>0.9987</u>

The results in Table 3 show that ViT has the lowest training loss (0.0296), but still has sub-optimal accuracy for some tumor classifications. This gap suggests that a small training loss may not necessarily lead to strong generalization. On the other hand, ResNet101 performs better accuracy on most of the labels, but with a slightly higher training loss. EfficientNet-B3 has the biggest training loss but shows better overall performance in the brain MRI classification job. We also include recent state-of-the-art designs for medical image analysis such as ConvNeXt and Swin Transformer in addition to the baseline models ViT and ResNet101. The results in Table 3 show that these models achieve a somewhat better accuracy than EfficientNet-B3. EfficientNet-B3 offers competitive performance and a more favorable trade-off between computational cost and accuracy, making it more suitable for real-world applications with limited resources. ConvNeXt achieved somewhat better accuracy than EfficientNet-B3 in our experiments, but we chose the latter as our main model due to a

better trade-off between predicted accuracy and computational cost. This is relevant for practical deployment in resource-constrained contexts. Additionally, the implementation of EfficientNet-B3 is well-established and widely used, which ensures the consistent optimization performance. EfficientNet-V2 was tested and showed competitive results, however the performance gain relative to the increased computational requirements was marginal compared to EfficientNet-B3 and hence EfficientNet-B3 was kept as the main model.

Table 4. Five-fold cross-validation results of EfficientNet-B3 on the brain MRI dataset

No	Metrics	Mean $\pm$ Std
1	Accuracy	0.9987229 $\pm$ 0.0017679
2	Precision	0.9987253 $\pm$ 0.0017625
3	Recall	0.9987229 $\pm$ 0.0017679
4	F1-score	0.9987228 $\pm$ 0.0017685

Moreover, in order to evaluate the robustness and generalization capabilities of the proposed EfficientNet-B3 model, we carried out a five-fold cross-validation experiment on the brain MRI dataset as shown in Table 4. The data set was randomly divided into five equal halves while preserving the class distribution by stratified sampling. In each cycle, 4 subsets (80%) were used for training and the other subset (20%) was used for validation. The process was repeated five times, each time using a different subset as validation data.

The overall performance was evaluated by averaging the results of the five folds. The standard deviation ($\pm$ Std) was also provided to evaluate the model stability. According to Table 4, EfficientNet-B3 has achieved the best performance in all folds with low standard deviation values for all measurements. This means that the model is accurate and robust and has low sensitivity to data segmentation.

The F1-score of 0.9987 ± 0.0018 means that the model is good in balancing precision and recall across subsets of data. The results confirm that EfficientNet-B3 has good generalization and low overfitting which is important for real-world medical applications.

3.4. Comparison of hospital recommendation system with neural network based collaborative filtering and K-NN

We compare our method neural network based collaborative filtering with K-NN to evaluate the accuracy of the hospital recommendation system. Specifically, the hospital data is first read, and then the CF model maps the tumor.id and hospital.id to embedding vectors, computes the dot product, normalizes the result with a sigmoid function, and multiplies by a scaling factor of 5. In contrast, the K-NN model uses a pivot table to calculate the Euclidean distance in order to find the k-nearest neighbors and recommend similar hospitals.

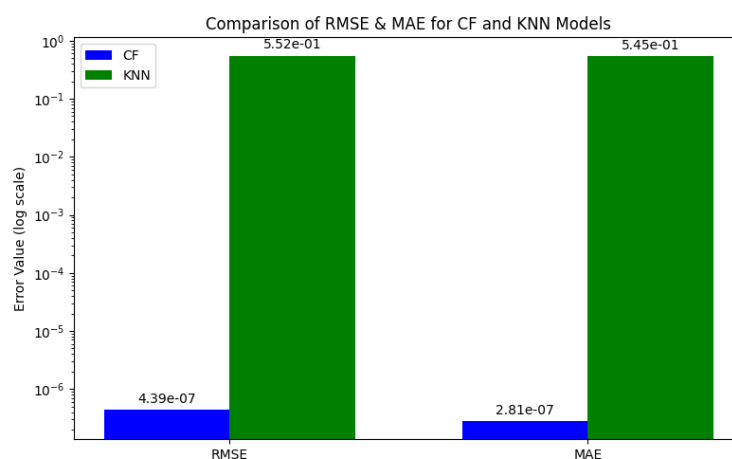


Figure 3. The comparison of RMSE and MAE for CF and KNN models

The results are presented in Fig. 3. It can be seen that the CF method achieves a very low error, whereas the KNN method exhibits noticeably higher errors. This significant discrepancy indicates that CF is

more stable and predicts ratings much more accurately. Although KNN initially shows promising results, its error increases substantially in the presence of noise. Consequently, Collaborative Filtering based on the embedding layer is confirmed to be a more reliable hospital recommendation solution in scenarios with distributed data and high accuracy requirements.

3.5. Discussion

The best accuracy is 85.38% for EfficientNet-B3, the second is 81.34% for ResNet-101 and the worst is 62.21% for ViT. The better performance of EfficientNet-B3 can be explained by its compound scaling method that balances the depth, width and resolution of the network in an ideal way, compared to traditional architectures. This allows to extract more complicated features while maintaining the computing efficiency. On the other hand, ResNet-101 uses deep residual connections to improve learning, but it does not scale as well for model performance. Vision Transformer (ViT) is a promising model in theory but often requires large datasets to obtain competitive performance. The lower score is either a sign of insufficient data or poor training because transformers usually struggle without sufficient pretraining. Therefore, the properly designed architecture of EfficientNet-B3 provides a competitive advantage in terms of accuracy and performance. Although Swin Transformer and ConvNeXt provide marginally better accuracy, they require much more computer effort and longer training time. In contrast, EfficientNet-B3 achieves the same performance with a more efficient architecture, which is useful for practical application in healthcare systems, especially in resource-limited contexts.

For the recommendation tasks, a neural network-based collaborative filtering approach significantly outperforms a K-Nearest Neighbors (K-NN) approach. The neural approach achieves a Mean Absolute Error (MAE) of $2.81\text{e-}07$ and a Root Mean Square Error (RMSE) of $4.39\text{e-}07$ as compared to the K-NN approach with an MAE of 0.545 and an RMSE of 0.552. This is mainly due to the ability of the neural network to learn complex non-linear interactions between users and items, with multiple layers and multiple embeddings, and hence to find more complex patterns in the data. On the other hand, K-NN is a memory-based, shallow approach which depends only on the proximity in the user or item similarity space, limiting its capability to generalize or deal with sparse data effectively. The accuracy and prediction performance of the neural model is better because it can learn from implicit feedback, user-item interactions and hidden information.

The combination of EfficientNet-B3 for MRI image classification and collaborative filtering for suggesting cancer treatment facilities provides a strong and holistic approach to personalized healthcare. EfficientNet-B3 with its improved accuracy and fine-tuned architecture efficiently captures small-scale patterns in MRI images for effective and timely cancer diagnosis. Collaborative filtering can be used to weigh the patient's preferences and treatment results once the patient is diagnosed and propose the most suitable facilities based on similar cases. This synergy can permit automated and accurate diagnosis, and personalized therapeutic recommendations, improving not only the clinical decision-making process but also the patient experience. The integration makes workflows easier for healthcare providers and gives patients personalized, data-driven support on their journey through treatment.

The drawback of this work was the MRI classification system was constructed from the public dataset rather than the local Vietnamese clinical data. This could affect the model's applicability to real settings of hospitals in Vietnam. In future work, we will focus on collecting and merging MRI data from Vietnamese hospitals, to strengthen the system validation and resilience.

The choice of EfficientNet-B3 should be viewed as a pragmatic design decision, rather than as a declaration that it is universally better than all present systems. Even if ConvNeXt or EfficientNet-V2 might achieve marginal gains, EfficientNet-B3 provides a good trade-off between accuracy, efficiency and ease of implementation for the integrated hospital recommendation system.

4. CONCLUSION

In this study, we set out to create an integrated solution combining deep learning based brain MRI classification with a hospital recommendation system, as outlined in the Introduction. Although the MRI dataset is not locally sourced, the hospital recommendation component is specifically designed for the Vietnamese healthcare system, providing practical value in real-world applications. The results confirm that our initial expectations were met: EfficientNet-B3 demonstrates strong classification performance for identifying different tumor types, while neural network based collaborative filtering provides stable and accurate hospital recommendations. These findings indicate that unifying high-accuracy diagnostic methods with a reliable recommendation engine can effectively support clinical decision-making and guide patients in choosing the most

suitable hospital. Furthermore, the system's scalability and precision can be enhanced in future work by incorporating larger or more diverse datasets, refining augmentation and recommendation algorithms, or utilizing distributed computing resources. By expanding on these results, subsequent studies can further optimize both the classification accuracy and recommendation quality, thereby advancing practical applications in healthcare.

REFERENCES

- [1] S. Gheewala, S. Xu, and S. Yeom, *n-depth survey: deep learning in recommender systems—exploring prediction and ranking models, datasets, feature analysis, and emerging trends* Neural Comput. Appl. 2025.
- [2] S. D and M. Krishnamurthy, "Recommendation system for frequent item sets using multi-objective chaotic optimization with convolutional BiLSTM model," *Appl. Intell.*, vol. 55, no. 7, p. 636, 2025, doi: 10.1007/s10489-025-06432-2.
- [3] Phuc Do, and Truong H.V. Phan, "Developing a BERT based triple classification model using knowledge graph embedding for question answering system," *Applied Intelligence*, vol. 52, p. 636-651, 2022, doi: <https://doi.org/10.1007/s10489-021-02460-w>.
- [4] Phuc Do, and Truong H.V. Phan, "Developing a Vietnamese Tourism Question Answering System Using Knowledge Graph and Deep Learning," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 5, p. 1-18, 2021, doi: <https://dl.acm.org/doi/abs/10.1145/3453651>.
- [5] V. H. Pham, A. T. Nguyen, B. T. Phung, and T. H.-V. Phan, "Developing a restaurant recommended system via the Vietnamese food image classification," *Int. J. Electr. Comput. Eng. IJECE*, vol. 14, no. 2, p. 1711, Apr. 2024, doi: 10.11591/ijece.v14i2.pp1711-1719.
- [6] V. B. Pham, P. C.-H. Nguyen, B. T. Phung, and T. H. V. Phan, "TourMapQA: using deep learning to develop a vietnam map-based tourism question answering system," *IAES Int. J. Artif. Intell. IJ-AI*, vol. 13, no. 3, p. 3203, 2024, doi: 10.11591/ijai.v13.i3.pp3203-3210.
- [7] T. H. V. Phan and P. Do, "BERT+vnKG: Using Deep Learning and Knowledge Graph to Improve Vietnamese Question Answering System," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, 2020, doi: 10.14569/IJACSA.2020.0110761.
- [8] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *J. Big Data*, vol. 9, no. 1, p. 59, 2022, doi: 10.1186/s40537-022-00592-5.
- [9] Y. Munir, Q. Umer, M. Faheem, S. Akram, and A. Jaffar, "Developer Recommendation and Team Formation in Collaborative Crowdsourcing Platforms," *IEEE Access*, pp. 1–1, 2025, doi: 10.1109/ACCESS.2025.3558557.
- [10] M. M. Reddy and P. Mohandas, "A User-based Collaborative Filtering Method to deal with Sparsity in Recommendation Systems by an unsupervised learning of Users' Hidden Preferences," in 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India: IEEE, Jun. 2022, pp. 14–20. doi: 10.1109/AIC55036.2022.9848963.
- [11] M. Suryani, R. Pangestu, and A. Pradana, "Optimizing Item-based Collaborative Filtering for Course Recommendation: a Case Study with Coursera Data," in 2024 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, Indonesia: IEEE, Dec. 2024, pp. 570–575. doi: 10.1109/ICITSI65188.2024.10929184.
- [12] K. E. Permana, "Comparison of User Based and Item Based Collaborative Filtering in Restaurant Recommendation System," *Math. Model. Eng. Probl.*, vol. 11, no. 7, pp. 1922–1928, Jul. 2024, doi: 10.18280/mmep.110723.
- [13] Z. Guan, "Collaborative Filtering and Sentiment Analysis: Basics to Build a Map Recommender System," in Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023), vol. 180, B. H. Ahmad, Ed., in Advances in Intelligent Systems Research, vol. 180, Dordrecht: Atlantis Press International BV, 2024, pp. 313–320. doi: 10.2991/978-94-6463-370-2_34.
- [14] S. Poudel and M. Bikdash, "Explaining the Performance of Collaborative Filtering Methods with Optimal Data Characteristics," *Int. J. Cybern. Inform.*, vol. 12, no. 2, pp. 287–295, 2023, doi: 10.5121/ijci.2023.120221.
- [15] A. Torkashvand, S. M. Jameii, and A. Reza, "Deep learning-based collaborative filtering recommender systems: a comprehensive and systematic review," *Neural Comput. Appl.*, vol. 35, no. 35, pp. 24783–24827, 2023, doi: 10.1007/s00521-023-08958-3.
- [16] X. Tian, "Content-based Filtering for Improving Movie Recommender System," in Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023), vol. 180, B. H. Ahmad, Ed., in Advances in Intelligent Systems Research, vol. 180, Dordrecht: Atlantis Press International BV, 2024, pp. 598–609. doi: 10.2991/978-94-6463-370-2_61.
- [17] K. Kosim and Reza Prihandi, "Recommendation System Algorithm Content-Based Filtering Method to Provide Drink Menu Recommendations," *J. Math. Instr. Soc. Res. Opin.*, vol. 2, no. 2, pp. 158–168, 2023, doi: 10.58421/misro.v2i2.130.
- [18] N. T. Anh and P. H. V. Truong, "Developing A Job Recommended System For Students In Information Technology Based On Learning Capability Meeting The Needs Of Employers," in Proceedings Of The 14th National Conference On Fundamental And Applied Information Technology Research, Hochiminh city, Vietnam, 2021. doi: 10.15625/vap.2021.0072.
- [19] F. Akram, T. Ahmad, and Mohd. Sadiq, "Recommendation systems-based software requirements elicitation process—a systematic literature review," *J. Eng. Appl. Sci.*, vol. 71, no. 1, p. 29, 2024, doi: 10.1186/s44147-024-00363-4.
- [20] A. Mulay, S. Sutar, J. Patel, A. Chhabria, and S. Mumbaikar, "Job Recommendation System Using Hybrid Filtering," *ITM Web Conf.*, vol. 44, p. 02002, 2022, doi: 10.1051/itmconf/20224402002.
- [21] Ruchika, M. Sharma, and S. A. Hossain, "Efficient Machine Learning Algorithms in Hybrid Filtering Based Recommendation System," *J. Inf. Technol. Manag.*, vol. 15, no. 3, 2023, doi: 10.22059/jitm.2023.93631.
- [22] B. R. Sri and T. S. Balakrishnan, "Classification of Pests in Agricultural Farms Using Convolutional Neural Network Compared to Artificial Neural Network," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India: IEEE, Jun. 2024, pp. 1–4. doi: 10.1109/ICCCNT61001.2024.10725825.
- [23] P. Shourie, V. Anand, and S. Gupta, "An Intelligent VGG-16 Framework for Brain Tumor Detection Using MRI-Scans," in 2023 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES), Tumakuru, India: IEEE, Jul. 2023, pp. 1–6. doi: 10.1109/ICSSES58299.2023.10200833.
- [24] G. H. T. Le, K. Nguyen-Tan, and H. X. Huynh, "Energy-Based Fairness Recommendation: A Novel Framework for Fairness Treatment in Recommender Systems," in 2025 5th Asia Conference on Information Engineering (ACIE), Phuket, Thailand: IEEE, Jan. 2025, pp. 125–128. doi: 10.1109/ACIE64499.2025.00027.
- [25] G. N, B. S, G. S, and M. S, "A Dual-Method Movie Recommendation System: Content and Collaborative," in 2025 Interna-

- tional Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India: IEEE, Feb. 2025, pp. 1953–1968. doi: 10.1109/ICEARS64219.2025.10941373.
- [26] S. Ahmadian, K. Berahmand, M. Rostami, S. Forouzandeh, P. Moradi, and M. Jalili, “Recommender Systems based on Non-negative Matrix Factorization: A Survey,” *IEEE Trans. Artif. Intell.*, pp. 1–21, 2025, doi: 10.1109/TAI.2025.3559053.
- [27] T. Amirifar, S. Lahmiri, and M. K. Zanjani, “An NLP-Deep Learning Approach for Product Rating Prediction Based on Online Reviews and Product Features,” *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 6, pp. 8156–8168, 2024, doi: 10.1109/TCSS.2023.3290558.
- [28] X. Huang, J. Wang, and J. Cui, “A Personalized Collaborative Filtering Recommendation System Based on Bi-Graph Embedding and Causal Reasoning,” *Entropy*, vol. 26, no. 5, p. 371, 2024, doi: 10.3390/e26050371.
- [29] N. S. Fatima et al., “Enhanced Facial Emotion Recognition Using Vision Transformer Models,” *J. Electr. Eng. Technol.*, vol. 20, no. 2, pp. 1143–1152, 2025, doi: 10.1007/s42835-024-02118-w.
- [30] G.-Y. Choi, N.-R. Kim, D.-Y. Yu, T. Lee, G. Lee, and H.-J. Hwang, “Transfer learning and data augmentation for glucose concentration prediction from colorimetric biosensor images,” *Microchim. Acta*, vol. 192, no. 5, p. 287, 2025, doi: 10.1007/s00604-025-07136-7.
- [31] W. Hu, J. Feng, and D. Yang, “An improved density peaks clustering algorithm using similarity assignment strategy with K-nearest neighbors,” *Clust. Comput.*, vol. 27, no. 9, pp. 12689–12706, 2024, doi: 10.1007/s10586-024-04592-3.
- [32] D. S. K. Karunasingha, “Root mean square error or mean absolute error? Use their ratio as well,” *Inf. Sci.*, vol. 585, pp. 609–629, Mar. 2022, doi: 10.1016/j.ins.2021.11.036.

BIOGRAPHIES OF AUTHORS



Anh Nhat Lam    holds a B.Sc. in information technology from Van Lang University, Ho Chi Minh City, Vietnam, in 2025. She is currently working as QC role at Momo corporation in Vietnam. Her research interests include Web programming, Artificial Intelligence, and Mobile. She can be contacted at email: nhathanh12@gmail.com.



Truong Ho-Viet Phan    was born in Ho Chi Minh City, Vietnam. He received the B.S. degree in informatics from Ton Duc Thang University, Ho Chi Minh City, in 2003, and the M.S. degree in computer science from the University of Information Technology, Viet Nam National University Ho Chi Minh, Vietnam, in 2013, where he is currently pursuing the Ph.D. degree in information technology. He has taught full-time with the Faculty of Information Technology, Van Lang University, since 2020. He is the author and co-author of nine ISI/Scopus journal articles. His research interests include social network analysis, text mining, chatbot, knowledge graphs, and question-answering systems. He can be contacted at email: truong.phv@vlu.edu.vn.



Bao The Phung    received the Ph.D. degrees in Information Technology from Irkutsk National Research Technical University (IrNITU), Irkutsk City, Russia, in 2014. He is currently the lecturer of Faculty of Information Technology, Ho Chi Minh City, University of Industry and Trade, Vietnam. His research interests include the Mathematical model, applications of artificial intelligence, deep learning, machine learning, data mining. He can be contacted at email: baopt@huit.edu.vn