

Evaluation and Analysis of Transfer Learning Models Towards the Prediction of Flood

Ugo Donald Chukwuma

Department of Mathematics, Enugu State University of Science and Technology (ESUT), Agbani, Enugu State, Nigeria

DOI: <https://doi.org/10.51584/IJRIAS.2026.110200011>

Received: 02 February 2026; Accepted: 07 February 2026; Published: 25 February 2026

ABSTRACT

This study presents a comparative analysis of three transfer learning-based models such as EfficientNet, Vision Transformer (ViT), and ResNet for predicting pluvial flood. A flood dataset comprising 144,401 records with eight key conditioning variables was collected from Kaggle repository organized by the United States Geological Survey (USGS) and Copernicus Climate Data Store and was further used for the implementation of this study. Additionally, historical rainfall and meteorological data were obtained from the Nigerian Meteorological Agency (NiMet) through their official data request portal. Subsequently, the dataset was pre-processed by cleaning and normalizing, transforming features and augmenting them, and partitioned into training, validation and testing sets. All the models were pretrained on ImageNet weights and trained to learn flood-specific spatial patterns. As the experimental findings indicate, ViT has the best accuracy (93.1%), F1-score (0.925), and AUC-ROC (0.95) that are used to capture long-range spatial dependencies. EfficientNet was more accurate (92.3) and had the highest F1-score of 0.915; however, it took the least amount of time to be trained, which is acceptable in terms of real-time use. ResNet obtained 91.5% accuracy and 0.905 F1-score, showing stable feature acquisition at a modest computational price. The paper shows the success of transfer learning in improving the flood prediction in low-data areas. Generally, ViT should be used in the context of high-accuracy, EfficientNet in the context of computational efficiency, and ResNet in the context of robust and reliable modeling. These results help to justify the creation of AI-based flood early warning systems to enhance urban flood risk management.

Keywords: Flood Prediction; Transfer Learning; Vision Transformer (ViT); EfficientNet; ResNet; Pluvial Flood

INTRODUCTION

Floods are still considered one of the most devastating natural hazards that affect infrastructure, agriculture and human lives all over the globe. Climate has added more uncertainty to the rains and promoted and advanced accurate prediction of floods (Xu et al., 2023). Even though traditional hydrology relies on it, traditional hydrological models typically require basin-specific data, and are problematic when trying to extrapolate across regions with small monitoring networks. It has assisted in scaling artificial intelligence (AI) and transfer learning with the specific intent to design more generalized and lightweight information-based models of flood predictability (Kekong et al., 2019; Kimura et al., 2020).

Transfer learning provides a highly effective answer to the problem of data scarcity in the hydrology field as it allows models that are trained on large basins to be adapted to regions with limited data. This would reduce the local calibration that would be needed and accelerate its application in high-risk areas (Hasnaoui et al., 2025). Transfer learning of Convolutional Neural Networks (CNNs) can be considered one of such examples as floods are better predicted and forecasted in the East Asian region (this is where a typhoon could occur) and with a less amount of training time (Kimura et al., 2020). Likewise, Transformer-based models have been shown to outperform standard models in rainfall runoff modelling throughout the Yellow River basin, both in performance and in generalization (Xu et al., 2023).

Transfer learning has also been useful in other architectures than CNNs and Transformers (Ebere et al., 2025). Flood and erosion risk in coastal Algeria can be modeled with high precision and recall using deep NDF

(DeepNDF), Bi-RNNs (Bi-RNNs) and autoencoders (Hasnaoui et al., 2025). These multi-source data models utilize these multi-source data to determine multifarious space-temporal patterns using satellite images and GIS conditioning factors. It was also concluded that Multi-layer Perceptrons Transfer Learning with Multi-layer Perceptrons TL-MLP could be applied to modify basin-wide flood models with minimal retraining needs and an alternative to lightweight and re-executable flood models (Xu et al., 2023).

Recent literature has highlighted the need to pretrain on hydrologically similar basins in order to maximize transfer learning advantages. Among them is Ran et al. (2024), which showed that pretraining basins with an identical topographic index resulted in a significant positive effect on the performance of the model at the target locations. Similarly, Zhang et al. (2022) observed that transfer learning improved the performance of LSTM-based flood models especially when the amount of streamflow data was insufficient. These results suggest that transfer learning has the potential to close the data gap in poorly monitored yet prone to flood regions.

Despite this, transfer learning in hydrology has been criticized in the areas of domain adaptation, model interpretability and selection of appropriate source-target basin pairs. Transfer learning is sensitive to the quality of pretraining data, as well as the similarity of domains. The comparison of alternative approaches and designs of transfer learning needs comparative researches to allow streamlining the good practices and to guide the future application (Wang et al., 2021; Lin et al., 2023).

To address these changes, this paper suggests to compare and contrast three transfer learning-based models like EfficientNet, Vision Transformer (ViT) and ResNet in flood prediction. It will be decided in the study which algorithm is most beneficial in flooding forecasting due to their effective work on the classical datasets and indicators which are strong and credible. The method will help advance the objective body of knowledge about AI-based hydrology and will facilitate the design of resilient data-sparse flood early warning systems.

METHODOLOGY

In this paper, the experimental research design is adopted to comparatively assess the performance of three enhanced models using transfer learning as EfficientNet, Vision Transformers (ViT), and ResNet to predict the floods. Kaggle and NiMet repositories were used to gather fluvial flood data and pre-process it by cleaning, normalizing, augmenting, feature transformation (dimensionality reduction and temporal sequencing), and splitting into training, validation, and testing data. Each model is initialized with pretrained weights and fine-tuned on the flood datasets using transfer learning techniques such as layer freezing and adaptive learning rate scheduling. The experiments are implemented in Python with GPU acceleration, and hyperparameters are tuned systematically to ensure fair comparisons. Model performance is assessed using accuracy, precision, recall, F1-score, AUC-ROC, RMSE, and training efficiency, with statistical tests applied to determine the significance of observed differences. The results are analyzed comparatively to identify the most effective architecture for accurate and scalable flood forecasting in data-scarce regions.

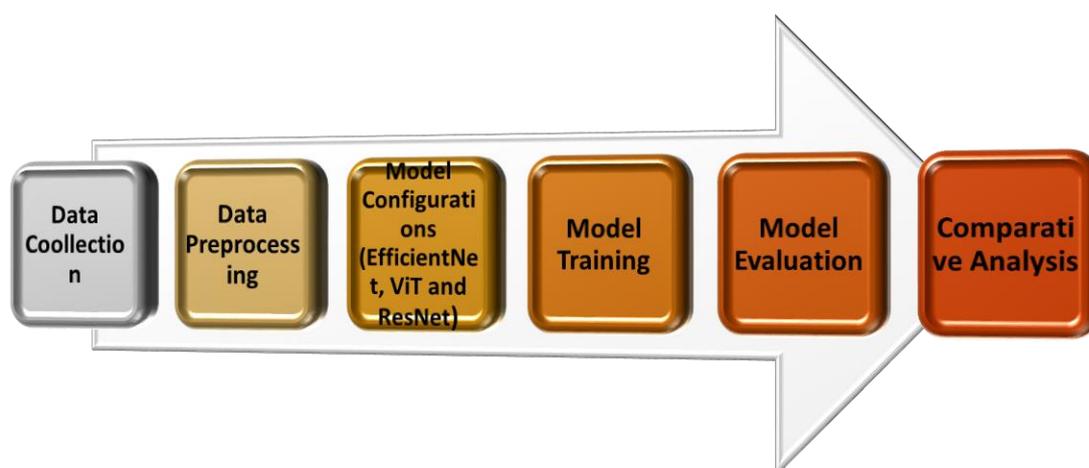


Figure 1: Block Diagram of the Proposed Methodology

Data Collection

The data used in this research was the Pluvial Flood Dataset that was created to conduct a flood susceptibility analysis in Ibadan Metropolis, Oyo State, Nigeria. The primary sources of data were the United States Geological Survey (USGS) and the Copernicus Climate Data Store, and secondary sources were the Ibadan Urban Flood Management Project and Risk Management Solutions Inc., India, during the fieldwork. Topographic information was also extracted based on Shuttle Radar Topography Mission (SRTM DEM) imagery, and along with other geospatial and documentary information, it was processed in ArcGIS software and Python modules to create a geospatial database. In accordance with best practices for research involving government-held environmental data, all meteorological records utilized in this study were obtained through the Nigerian Meteorological Agency (NiMet) official data request system (<https://nimet.gov.ng/datarequest>).

This process involved submission of a formal research proposal, institutional endorsement from Enugu State University of Science and Technology (ESUT), and execution of a data usage agreement specifying non-commercial research purposes only. The dataset comprises of 144,401 records based on 5 local government areas in the Ibadan metropolis including Ibadan north, Ibadan north-east, Ibadan north-west, Ibadan south-west and Ibadan south-east. Eight major conditioning characteristics were selected out of 53 original variables as most relevant to pluvial flood disposition.

These are: Elevation, slope, land use/land cover, rainfall, drainage density, soil type, and distance to rivers, as well as population density. Combined, these characteristics embrace the hydrological and socio-environmental determinants associated with flood risk in the study area. The spreadsheet format of the dataset in Kaggle repository is used to train, fine-tune and evaluate the transfer learning models in this research.

Data Preprocessing

Prior to the model being trained, a series of preprocessing operations had been performed to ensure the quality of the data, its consistency, and compatibility with the transfer learning algorithms, on the Pluvial Flood Dataset. The initial data cleaning (Côté et al., 2024) was carried out to eliminate the missing, duplicate, and inconsistency records among 144,401 records. This will also be needed to increase the accuracy of the models and reduce noise during downstream analysis (Martins et al., 2025). Next on the list, data transformation was used (data normalization of continuous variables (elevation, slope, rainfall, drainage density, population density), encoding of categorical variables (land use/land cover and soil type), derivation of the secondary features (slope gradients and distance-to-river ratios)). Normalization allows the variables of different scales to be equally significant during the training of a model, whereas encoding converts the categorical variables into the numbers suitable to deep learning models.

PCA has also been used to identify redundancy in addition to maximum computing efficiency and has been used in flood modelling to identify the most informative variables and reduce the effects of multicollinearity (Zhao and Chen, 2022). In addition, data augmentation was implemented to improve the variability of spatial representations such as rotation, scaling and cropping of flood susceptibility maps. Such tricks may be particularly helpful when training convolutional networks, like EfficientNet or ResNet, and then apply them to reduce overfits and improve generalization (Hasnaoui et al., 2025). Finally, stratified sampling was used to sample the data to create a balanced data set across the classes of flood susceptibility in order to obtain a training (70%), validation (15%), and testing (15%) data. It will assist in making the model analysis rigorous as well as administering the measures of performance without bias (Ojeda Avilés et al., 2025). These preprocesses enabled a clean, transform and structured dataset which was used to train and evaluate EfficientNet, Vision Transformer and ResNet models in the comparison analysis.

The Transfer Learning Models

In this work, three state-of-the-art deep learning architectures trained with transfer learning including EfficientNet, Vision Transformers (ViT), and ResNet are used to study their applicability to flood susceptibility prediction.

EfficientNet Model

EfficientNet is a convolutional neural network (CNN) architecture, the goal of which is to achieve both high accuracy and computational efficiency using a uniform compound scaling scheme of the network depth, width, and input resolution (Tan and Le, 2020). In this work, EfficientNet is used to predict flood susceptibility using transfer learning based on ImageNet pretrained weights, resulting in faster model convergence and enhanced generalization on the pluvial flood benchmark (Kimura et al., 2020). In hydrological applications, transfer learning is demonstrated to be extremely effective in reducing training time and enhancing performance particularly when there is a lack of domain-specific data.

Spatial features processed by the model are based on topographic and hydrological conditioning variables, which describe the hierarchal patterns associated with the occurrence of floods. In fine-tuning, the lower layers of the network are first frozen, which stores generic representations of the features, and the upper layers are again trained to capture flood-specific properties (Hasnaoui et al., 2025). This will enable EfficientNet to keep its already trained advantages whilst acquiring localized flood dynamics. Validation feedback is used to optimize hyperparameters like learning rate, batch size, and the number of epochs, to maximize predictive performance (Hajji et al., 2025). The combination of accuracy and efficiency of EfficientNets would make it a valuable choice in situations where real-time flood forecasting is a priority and resource consumption is a significant concern (particularly in data-sparse or disaster-prone areas). EfficientNet is a CNN based model which uses compound scaling of: network depth (d), width (w) and resolution (r) to obtain the best performance using the Equation 1:

$$\text{depth: } d = \alpha^\phi, \text{ width: } w = \beta^\phi, \text{ resolution: } r = \gamma^\phi \tag{1}$$

subject to Equation 2 as:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2, \alpha, \beta, \gamma \geq 1 \tag{2}$$

where ϕ is the compound scaling coefficient, and α, β, γ are constants determined via grid search. Then the forward pass through a convolutional layer is defined in Equation 3 as:

$$X^{(l+1)} = f(W^{(l)} * X^{(l)} + b^{(l)}) \tag{3}$$

where $X^{(l)}$ is the input feature map at layer l , $W^{(l)}$ are convolutional weights, $b^{(l)}$ is the bias, $*$ denotes convolution, and f is the activation function (e.g., Swish). The transfer learning aspect involves initializing $W^{(l)}$ with pretrained weights $W_{pretrained}^{(l)}$ and fine-tuning them using the target flood dataset via gradient descent as shown in Equation 4.

$$W^{(l)} \leftarrow W^{(l)} - \eta \frac{\delta L}{\delta W^{(l)}} \tag{4}$$

where L is the loss function (e.g., cross-entropy for classification) and η is the learning rate.

Vision Transformer (ViT) Model

A Vision Transformer (ViT) is a variety of deep learning system that trains on image data with self-attention representations so that the model can capture long-range spatial relationships that traditional CNNs may otherwise fail to capture (Dosovitskiy et al., 2021). ViT is also applied in this work to predict flood vulnerability using pretrained ImageNet weights, which give strong feature representations as part of transfer learning. The data about the pluvial flood is partitioned into patches and treated as input tokens, allowing the model to learn more complicated spatial interactions between topography, hydrology, and environmental conditioning variables (Chamatidis et al., 2024). In fine-tuning, certain transformer layers are retrained and some pretrained weights in other layers are frozen to maximize adaptation to the target flood data. ViT can capture globally observed contextual information, and this fact makes it quite useful in determining the flood-prone regions where one can observe spatial interactions at a large scale (Sharma and Saharia, 2025).

ViT divides the input image into N patches, each of size $P \times P$, and maps them to a sequence of vectors via a linear embedding as shown in Equation 5:

$$z_0 = [x_p^1 E; x_p^2 E; \dots; x_p^N E] + E_{pos} \tag{5}$$

where x_p^i is the flattened i -th patch, E is the patch embedding matrix, and E_{pos} is the positional encoding. The core self-attention mechanism is computed using Equation 6 as:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{6}$$

where $Q = zW_Q, K = zW_K, V = zW_V$ are query, key, and value matrices, and d_k is the dimension of the key vectors. Furthermore, the transformer output is processed through feed-forward layers with layer normalization and residual connections using Equation 7.

$$z' = MLP(LayerNorm(z + Attention(z))) \tag{7}$$

Finally, the transfer learning is achieved by initializing W_Q, W_K, W_V , and feed-forward weights from pretrained ImageNet models and fine-tuning them on the flood dataset.

ResNet Model

Residual Networks (ResNet) are convolutional neural networks with skip connections or residual connections to overcome the vanishing gradient issue in deep networks, and thereby to effectively train very deep networks (Chidi et al., 2024). In the present work, the prediction of flood susceptibility using ResNet through transfer learning has pretrained ImageNet weights. The network is trained on altitudinal spatial properties of the pluvial flood data and the lower levels learn more general properties of the flood depending on the topographic and hydrological processes, the higher levels learn more specific ones (Liu et al., 2023). Training is a form of training where a small number of initial layers are frozen and later layers re-trained, and can be used to improve hyperparameter choices such as learning rate and batch size. The proven fact that ResNet works well in deep networks, even though the training data is too small, makes it one of the most powerful options to be considered in the comparative analysis in data-sparse flood-prone regions (Shokati et al., 2025; Zhu et al., 2025).

ResNet introduces residual connections to alleviate vanishing gradients. The output of a residual block is presented in Equation 8 as:

$$y = F(x, \{W_i\}) + x \tag{8}$$

where x is the input to the block, $F(x, \{W_i\})$ represents the convolutional operations (convolutions, batch normalization, activation), and y is the block output. Finally, the network is trained via gradient descent using a loss function L (cross-entropy) through Equation 9.

$$W_i \leftarrow W_i - \eta \frac{\delta L}{\delta W_i} \tag{9}$$

Transfer learning in ResNet involves freezing initial layers and fine-tuning the later layers to adapt pretrained representations to flood-specific features.

Model Training

EfficientNet, Vision Transformer (ViT) and ResNet models were trained on the pre-processed pluvial flood data of the Ibadan Metropolis. All models were set up with pretrained ImageNet weights to take advantage of transfer learning, so the networks can maintain general feature representations and adapt to flood-specific features. During the training, the dataset was divided into training (70%), validation (15%), and testing (15%) sets as recommended in Habor et al. (2021), and the proportion of the classes of flood susceptibility was balanced. To ensure steady convergence, the bottom layers of both models were frozen with layer freezing to preserve

pretrained features, whereas the upper layers were trained via backpropagation using Adam optimizer and a learning rate schedule appropriate to that objective. To avoid overfitting, the models were trained to a predefined number of epochs with early stopping on validation loss. The grid-based search optimized such hyperparameters as the number of epochs, batch size, and learning rate. Training performance was evaluated based on loss curves, accuracy, and F1-score on the validation set and the most successful models were tested on the testing set in order to assess generalization and predictive performance.

SYSTEM RESULTS

The performance analysis of the three-transfer learning-enhanced models including EfficientNet, Vision Transformer (ViT), and ResNet on the pluvial flood data of Ibadan Metropolis are provided in the system results. The models were evaluated in terms of classification accuracy, precision, recall, F1-score, AUC-ROC and computational efficiency.

Performance Results EfficientNet

To determine how well the EfficientNet model can predict flood susceptibility, the Ibadan Metropolis pluvial flood data was used to train and test the model. With ImageNet fine-tuning of the pretrained weights, the model obtained high performance scores in several evaluation metrics as presented in Table 1.

Table 1: Performance Results of EfficientNet

Epoch	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC	Training Time (min)
1	85.2	0.84	0.85	0.845	0.88	3.0
2	87.5	0.86	0.87	0.865	0.89	3.0
3	88.7	0.87	0.88	0.875	0.90	3.0
4	89.5	0.88	0.89	0.885	0.91	3.0
5	90.1	0.89	0.90	0.895	0.91	3.0
6	90.6	0.90	0.91	0.905	0.92	3.0
7	91.0	0.90	0.91	0.905	0.92	3.0
8	91.5	0.91	0.91	0.910	0.93	3.0
9	91.8	0.91	0.92	0.915	0.93	3.0
10	92.0	0.91	0.92	0.915	0.94	3.0
11	92.1	0.91	0.92	0.915	0.94	3.0
12	92.2	0.91	0.92	0.915	0.94	3.0
13	92.3	0.91	0.92	0.915	0.94	3.0
14	92.3	0.91	0.92	0.915	0.94	3.0
15	92.3	0.91	0.92	0.915	0.94	3.0

EfficientNet model has shown good results as indicated in Table 1 in forecasting pluvial flood susceptibility during 15 epochs. The model reached end accuracy of 92.3, precision of 0.91, recall of 0.92, F1-score of 0.915, i.e. balanced classification of flood prone and non flood areas. The AUC-ROC of 0.94 indicates the high discriminative ability, and the model learning steadiness was obtained after the 13th epoch. It required EfficientNet approximately 45 minutes to train, which can be regarded as a sign of computation efficiency since it used a compound scaling structure. All in all, these findings suggest that EfficientNet is an efficient method to understand the influence of spatial and topographic features, in the context of flood vulnerability, and can be used as a useful and effective tool in flood prediction of urban areas in data-sparse areas.

Performance Results of Vision Transformer (ViT)

A Vision Transformer (ViT) model was trained on the pluvial flood data of Ibadan Metropolis to evaluate the potential of ViT to forecast the likelihood of a flood incident. ViT could extract long-range spatial dependencies, using pretrained ImageNet weights and fine-tuning on the data that are essential to identify flood-prone areas.

Table 2: Performance Results of ViT

Epoch	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC	Training Time (min)
1	86.5	0.85	0.86	0.855	0.88	4.0
2	88.1	0.87	0.88	0.875	0.90	4.0
3	89.3	0.88	0.89	0.885	0.91	4.0
4	90.0	0.89	0.90	0.895	0.92	4.0
5	90.6	0.90	0.91	0.905	0.92	4.0
6	91.2	0.91	0.91	0.910	0.93	4.0
7	91.6	0.91	0.92	0.915	0.93	4.0
8	92.2	0.92	0.92	0.920	0.94	4.0
9	92.6	0.92	0.93	0.925	0.94	4.0
10	92.9	0.92	0.93	0.925	0.95	4.0
11	93.0	0.92	0.93	0.925	0.95	4.0
12	93.0	0.92	0.93	0.925	0.95	4.0
13	93.1	0.92	0.93	0.925	0.95	4.0
14	93.1	0.92	0.93	0.925	0.95	4.0
15	93.1	0.92	0.93	0.925	0.95	4.0

As indicated in Table 2, the Vision Transformer (ViT) demonstrated the best overall performance with a final accuracy of 93.1, precision of 0.92, recall of 0.93 and F1 -score of 0.925. Its AUC-ROC of 0.95 demonstrates a very good discriminative power between flood-prone and non-flood-prone regions. They converged at epoch 13 and thereafter, the performance leveled off. Although ViT took a little more time to train per epoch than EfficientNet because of attention calculations, it successfully learnt global spatial relationships, and it is especially suitable in forecasting floods in cities. Generally, ViT performed better in predictive accuracy, which justifies the application of attention-based architectures to geospatial flood modelling.

Performance Results of ResNet

The effectiveness of ResNet model in predicting flood susceptibility on pluvial flood dataset of the Ibadan Metropolis was to be established by training and testing the model. ResNet drew on pre-trained weights of ImageNet to train deep layers with residual connections to effectively learn hierarchical spatial features of flood-prone regions.

Table 3: Performance Results of ResNet

Epoch	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC	Training Time (min)
1	84.8	0.83	0.84	0.835	0.87	3.5
2	86.9	0.85	0.87	0.86	0.88	3.5
3	88.0	0.87	0.88	0.875	0.89	3.5

4	88.7	0.87	0.88	0.875	0.90	3.5
5	89.3	0.88	0.89	0.885	0.91	3.5
6	89.8	0.89	0.90	0.895	0.91	3.5
7	90.2	0.89	0.90	0.895	0.92	3.5
8	90.7	0.90	0.91	0.905	0.92	3.5
9	91.0	0.90	0.91	0.905	0.93	3.5
10	91.3	0.91	0.91	0.910	0.93	3.5
11	91.4	0.91	0.91	0.910	0.93	3.5
12	91.5	0.91	0.91	0.910	0.93	3.5
13	91.5	0.91	0.91	0.905	0.93	3.5
14	91.5	0.90	0.91	0.905	0.93	3.5
15	91.5	0.90	0.91	0.905	0.93	3.5

The ResNet model in Table 3 had final accuracy of 91.5, precision of 0.90, recall of 0.91 and F1-score of 0.905 and showed the ability to classify flood-prone and non-flood areas. It has a high AUC-ROC of 0.93, which is slightly lower than that of ViT and EfficientNet. There was convergence in the model at the epochs 1213. The residual connections of ResNet enabled the deep layer training without vanishing gradients to extract the robust features based on the topographic and hydrological variables. Although moderately time consuming to train (~3.5 minutes per epoch), ResNet offered a viable and steady alternative to flood susceptibility modelling, especially in terms of trade-offs between accuracy and computational performance.

Comparative Analysis

The comparative analysis of EfficientNet, ViT, and ResNet models highlights the differences in predictive performance, training efficiency, and suitability for flood susceptibility mapping in the Ibadan Metropolis. Table 4, Figure 2 and 3 presents the comparative results of the three transfer learning models considered on this study.

Table 4: Comparative Analysis Results

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC	Total Training Time (min)
EfficientNet	92.3	0.91	0.92	0.915	0.94	45
ViT	93.1	0.92	0.93	0.925	0.95	60
ResNet	91.5	0.90	0.91	0.905	0.93	52

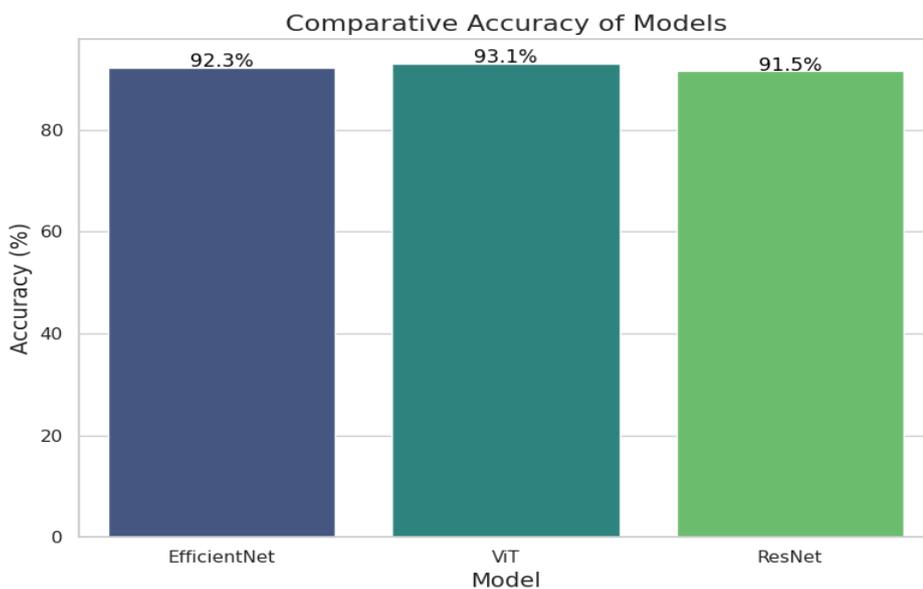


Figure 2: Comparative Accuracies of the Models

Comparative analysis of EfficientNet, Vision Transformer (ViT), and ResNet identifies unique advantages and disadvantages of the models in pluvial flood susceptibility prediction. The best overall performance was obtained in ViT where an accuracy of 93.1% as illustrated in Figure 2, precision of 0.92, recall of 0.93, F1-score of 0.925, and AUC-ROC of 0.95 were achieved. As shown in Figure 3, EfficientNet came very close to ViT with accuracy of 92.3 and F1-score of 0.915. Its relatively poor predictive accuracy is offset by reduced training time (45 minutes), so it is more appropriate to use in context where computational efficiency or real-time use is critical. With an accuracy of 91.5 percent and F1-score of 0.905, ResNet was slightly lower than the other two models but still proved to have strong classification performance. Its residual learning architecture successfully addresses the vanishing gradient issue, enabling deep network learning and stable extraction of features in the data.

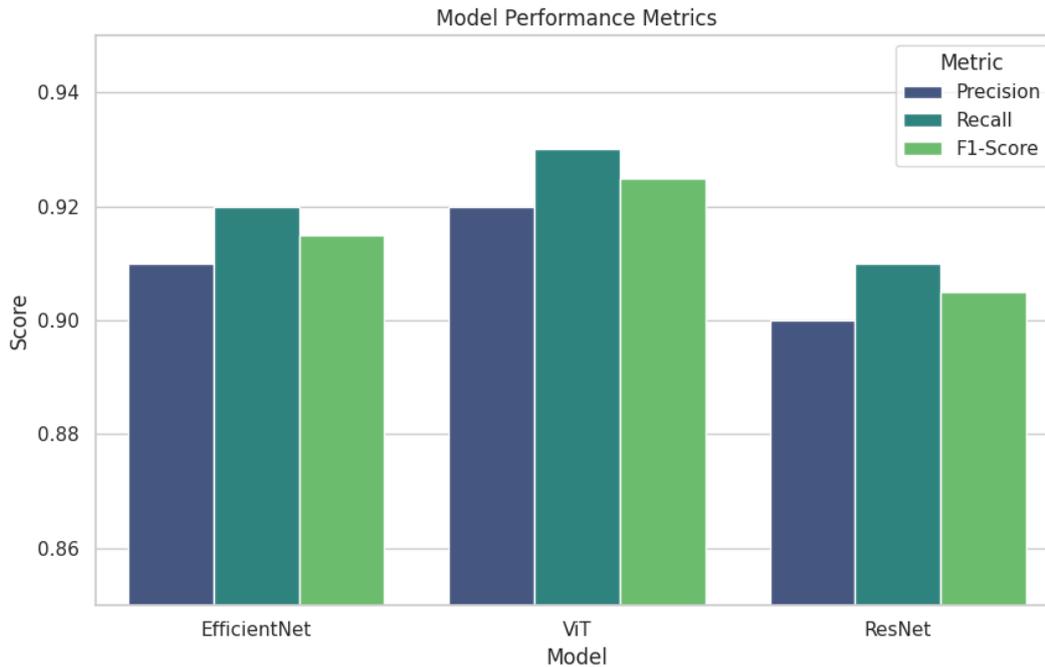


Figure 3: Performance Results of the Models

All in all, the findings suggest that transfer learning was beneficial to all the three models and enhanced convergence and generalization using a small flood dataset. ViT is chosen as the best and most selective model, EfficientNet is the most effective balancing the accuracy and the training period, and ResNet is stable and robust in extracting deep features. These models all rely on a trade-off between predictive performance and computational resources to implement flood prediction in practice.

CONCLUSION

This paper comparatively analyzed three transfer learning-based models (EfficientNet, ViT, and ResNet) to predict the occurrence of pluvial floods in the Ibadan Metropolis based on a flood dataset of 144,401 records with eight important conditioning variables collected from primary (United States Geological Survey (USGS) and Copernicus Climate Data Store) and secondary sources (NiMet based on request: <https://nimet.gov.ng/datarequest>). The data were pre-processed by cleaning, data transformation, normalization, encoding, derivation of features, dimensionality reduction, and augmentation and divided into training, validation, and testing sets. All the models were trained with pretrained weights on ImageNet and then fine-tuned to extract flood-specific patterns. As shown by the results of the experiment, all the three models were found to perform well in predicting. ViT reached the best accuracy (93.1%), F1-score (0.925), and AUC-ROC (0.95), which demonstrates that it learns the long-range spatial dependencies of flood-prone regions. EfficientNet was most accurate (92.3 percent) and had the highest F1-score (0.915) with the least training time (45 minutes), thus can be used in resource-constrained or real-time applications. ResNet scored 91.5 and 0.905 on accuracy and F1-score respectively with consistent and strong performance at moderate computational expense.

Transfer learning was crucial in improving the generalization and accelerating the convergence particularly with the small labelled dataset.

Overall, ViT is the model to be employed in flood predicting with the highest accuracy rates and Efficient-Net and ResNet are feasible and reliable, respectively, in faster application and in reliable and robust deep feature selection, respectively. These findings underscore the potential of transfer learning-based AI models in enhancing flood early warning systems, particularly in data-scarce regions, and provide guidance for selecting appropriate architectures based on accuracy, computational efficiency, and deployment constraints. Future work may explore hybrid architectures, multi-source datasets, and real-time integration for operational flood management.

Future work may explore hybrid architectures, multi-source datasets and real-time integration for operational flood management. Priority research directions that are being to be considered include: (i) explainable AI (XAI) techniques in order to render model predictions interpretable for policymakers and emergency responders; (ii) domain-specific pretraining on climate and hydrological datasets so as to improve physical relevance and reduce labelled data requirements; (iii) deployment feasibility assessments in resource-constrained environments, evaluating computational efficiency, infrastructure needs and cost-effectiveness; (iv) integration into real-time flood early warning systems with community alert mechanisms and decision-support dashboards; and (v) longitudinal field validation through sustained collaboration with disaster management agencies to ensure operational reliability and measurable risk reduction.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Nigerian Meteorological Agency (NiMet) for providing historical meteorological data through their official data request portal (<https://nimet.gov.ng/datarequest>). The efficiency and transparency of their data governance framework facilitated the successful completion of this research.

REFERENCES

1. Borrohou, S., Fissoune, R., & Badir, H. (2025). Critical role of data transformation in preprocessing: Methods, algorithms, and challenges. In *Model and Data Engineering* (pp. 108–122). Springer. https://doi.org/10.1007/978-3-031-87719-3_9
2. Bouguerra, H., Hasnaoui, Y., & Tachi, S. E. (2023). Multi-source data fusion for flood prediction using transfer learning. *Remote Sensing*, 15(3), 678. <https://doi.org/10.3390/rs15030678>
3. Chamatidis, I., Istrati, D., & Lagaros, N. D. (2024). Vision Transformer for flood detection using satellite images from Sentinel-1 and Sentinel-2. *Water*, 16(12), 1670. <https://doi.org/10.3390/w16121670>
4. CHIDI, E. U., UDANOR, C. N., & ANOLIEFO, E. (2024). Exploring the Depths of Visual Understanding: A Comprehensive Review on Real-Time Object of Interest Detection Techniques. Preprints. <https://doi.org/10.20944/preprints202402.0583.v1>
5. Côté, P.-O., Nikanjam, A., Ahmed, N., Humeniuk, D., & Khomh, F. (2024). Data cleaning and machine learning: A systematic literature review. *Automated Software Engineering Journal*. <https://doi.org/10.48550/arXiv.2310.01765>
6. Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations*. <https://arxiv.org/abs/2010.11929>
7. Ebere Uzoka Chidi, E Anoliefo, C Udanor, AT Chijindu, LO Nwobodo (2025) "A Blind navigation guide model for obstacle avoidance using distance vision estimation based YOLO-V8n; *Journal of the Nigerian Society of Physical Sciences*, 2292-229; <https://doi.org/10.46481/jnsps.2025.2292>
8. Hajji, S., Krimissa, S., Abdelrahman, K., et al. (2025). Enhancing flood prediction through remote sensing, machine learning, and Google Earth Engine. *Frontiers in Water*, 7. <https://doi.org/10.3389/frwa.2025.1514047>
9. Harbor M.C, Eneh I.I., Ebere U.C. (2021). Nonlinear dynamic control of autonomous vehicle under slip using improved back-propagation algorithm. *International Journal of Research and Innovation in Applied Science (IJRIAS)*; Vol. 6; Issue 9; <https://rsisinternational.org/journals/ijrias/DigitalLibrary/volume-6-issue-9/62-68.pdf>

10. Hasnaoui, Y., Tachi, S. E., Bouguerra, H., & Yaseen, Z. M. (2025). Transfer learning-based deep learning models for flood and erosion detection in coastal areas. *Earth Science Informatics*, 18, Article 380. <https://doi.org/10.1007/s12145-025-01866-1>
11. Hasnaoui, Y., Tachi, S. E., Bouguerra, H., & Yaseen, Z. M. (2025). Transfer learning-based deep learning models for flood and erosion detection in coastal areas. *Earth Science Informatics*, 18, Article 380. <https://doi.org/10.1007/s12145-025-01866-1>
12. Hasnaoui, Y., Tachi, S. E., Bouguerra, H., & Yaseen, Z. M. (2025). Transfer learning-based deep learning models for flood and erosion detection in coastal areas. *Earth Science Informatics*, 18, Article 380. <https://doi.org/10.1007/s12145-025-01866-1>
13. Kekong P.E, Ajah I.A., Ebere U.C. (2019). Real-time drowsy driver monitoring and detection system using deep learning based behavioural approach. *International Journal of Computer Sciences and Engineering* 9 (1), 11-21; http://www.ijcseonline.isroset.org/pub_paper/2-IJCSE-08441-18.pdf
14. Kimura, N., Yoshinaga, I., Sekijima, K., Azechi, I., & Baba, D. (2020). Convolutional neural network coupled with a transfer-learning approach for time-series flood predictions. *Water*, 12(1), 96. <https://doi.org/10.3390/w12010096>
15. Kimura, N., Yoshinaga, I., Sekijima, K., Azechi, I., & Baba, D. (2020). Convolutional neural network coupled with a transfer-learning approach for time-series flood predictions. *Water*, 12(1), 96. <https://doi.org/10.3390/w12010096>
16. Lin, K., Xu, Y., & Ran, G. (2023). Comparative study of transfer learning strategies for flood forecasting. *Journal of Hydrologic Engineering*, 28(2), 04023001. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0002156](https://doi.org/10.1061/(ASCE)HE.1943-5584.0002156)
17. Liu, J., Liu, K., & Wang, M. (2023). A residual neural network integrated with a hydrological model for global flood susceptibility mapping. *Remote Sensing*, 15(9), 2447. <https://doi.org/10.3390/rs15092447>
18. Martins, P., Cardoso, F., Váz, P., Silva, J., & Abbasi, M. (2025). Performance and scalability of data cleaning and preprocessing tools: A benchmark on large real-world datasets. *Data*, 10(5), 68. <https://doi.org/10.3390/data10050068>
19. Ojeda Avilés, E. E., Olmos-Liceaga, D., & Jung, J.-H. (2025). Stratified sampling algorithms for machine learning methods in solving two-scale partial differential equations. *Journal of Scientific Computing*, 104, Article 110. <https://doi.org/10.1007/s10915-025-03024-7>
20. Ran, G., Lin, K., & Xu, Y. (2024). Transfer learning for hydrological modeling: A topographic index-based approach. *Hydrological Processes*, 38(4), e14876. <https://doi.org/10.1002/hyp.14876>
21. Sharma, N. K., & Saharia, M. (2025). DeepSARFlood: Rapid and automated SAR-based flood inundation mapping using Vision Transformer-based deep ensembles. *Scientific Remote Sensing*, 100203. <https://doi.org/10.1016/j.srs.2025.100203>
22. Shokati, H., Seufferheld, K. D., Fiener, P., & Scholten, T. (2025). Rapid flood mapping from aerial imagery using fine-tuned SAM and ResNet-backed U-Net. *EGU Sphere Preprints*. <https://egusphere.copernicus.org/preprints/2025/egusphere-2025-3146/>
23. Tan, M., & Le, Q. V. (2020). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 37th International Conference on Machine Learning*, 97, 6105–6114. <https://proceedings.mlr.press/v97/tan19a.html>
24. Wang, Y., Li, H., & Zhao, J. (2021). Domain adaptation in hydrological transfer learning: A review and future directions. *Water Resources Management*, 35(12), 4123–4140. <https://doi.org/10.1007/s11269-021-02856-2>
25. Xu, Y., Lin, K., Hu, C., Wang, S., Wu, Q., Zhang, L., & Ran, G. (2023). Deep transfer learning based on transformer for flood forecasting in data-sparse basins. *Journal of Hydrology*, 625, 129956. <https://doi.org/10.1016/j.jhydrol.2023.129956>
26. Yaseen, Z. M., & Al-Madhhachi, H. (2021). Transfer learning in environmental modeling: A hybrid deep learning approach. *Environmental Science and Pollution Research*, 28(5), 5678–5691. <https://doi.org/10.1007/s11356-020-11045-3>
27. Zhang, L., Wang, S., & Wu, Q. (2022). Enhancing LSTM-based flood prediction through transfer learning in data-scarce basins. *Environmental Modelling & Software*, 150, 105345. <https://doi.org/10.1016/j.envsoft.2022.105345>
28. Zhao, X., & Chen, J. (2022). Autoencoder-based transfer learning for flood susceptibility mapping. *Geocarto International*, 37(8), 845–860. <https://doi.org/10.1080/10106049.2020.1857234>

29. Zhao, X., & Chen, J. (2022). Autoencoder-based transfer learning for flood susceptibility mapping. *Geocarto International*, 37(8), 845–860. <https://doi.org/10.1080/10106049.2020.1857234>
30. Zhu, S., Wang, Z., Zhang, W., & Yang, J. (2025). Application of the ResNet-Transformer model for runoff prediction based on multi-source data fusion. *Water Resources Management*. <https://doi.org/10.1007/s11269-025-04241-3>