OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 2 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

SWARM INTELLIGENCE-BASED FEATURE SELECTION FOR SKIN CANCER DATA ANALYSIS USING CONVOLUTIONAL NEURAL NETWORKS

Kiran Manzoor
Department of Computer Science, NFCIET, Multan, Pakistan

Yasir Aziz
Department of Computer Engineering, BZU, Multan, Pakistan

Muhammad Kamran Abid
Department of Computer Science, NFCIET, Multan, Pakistan

Naeem Aslam
Department of Computer Science, NFCIET, Multan, Pakistan

Muhammad Fuzail
Department of Computer Science, NFCIET, Multan, Pakistan

Nasir Umer
Department of Computer Science, NFCIET, Multan, Pakistan

[*]***Corresponding author:*** *engr.yasiraziz@bzu.edu.pk*
DOI: *https://doi.org/10.71146/kjmr283*

Article Info

Abstract

In this paper, a comparative analysis of three widely used deep learning models to perform data classification tasks with and without a Swarm Intelligence (SI) based Feature Selection is presented. The study explores the performance of each model in the context of types of data it can handle and the effect of SI-based feature selection on its performance. Traditionally used for image-based tasks, CNNs outperformed RNNs and LSTMs with a baseline accuracy of 90%, capitalizing on their strength to learn spatial hierarchies in image data. The feasibility of SI for feature selection was demonstrated on RNNs in retrieving data time series (reaching 85% accuracy, which was 5% higher than an 80% baseline accuracy achieved without feature selection). Additionally, we investigated the generation and the execution of data collection, tasks which we encoded in the limited capability decision making module and subsequently plan to contextualize with appropriate application knowledge. SI based feature selection through Particle Swarm Optimization (PSO) has enhanced the performance of all model by selecting the most paneled features and brought the dimensionality of the input data on much lesser side. For CNNs we demonstrated a 4% gain in accuracy from 90% to 94%. In terms of RNNs and LSTMs the improvement was a bit smaller (but not by much), so that the accuracies rose from 80% to 85% for RNNs, and from 85% to 88% for LSTMs. The results demonstrate the significance of feature selection in deep learning models with high dimensional datasets because it not only improves accuracy but also decreases model deviation and computational expense. The study concludes generally that CNN models are better for image-based tasks while RNN models and LSTMs are better for sequential data. The integration of Swarm Intelligence-based feature selection worked out to be effective for all the models, promising the potential to increase the deep learning performance in practical applications.

Keywords:
Skin Cancer, CNN, Swarm Intelligence, LSTM, RNN

Introduction

Skin cancer is by far the most common cancer in the world with increasing incidence and can be a serious disease with important public health implications. Especially the leading types of skin cancers, non-melanoma and melanoma, include basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), respectively. Moles are less likely to turn into cancer, but if those cancers do occur, they tend to be more serious, since moles may spread more rapidly to other parts of the body[1]. Conversely, melanoma is less common than BCC and SCC but is the deadlier of the two. However, timely cancer detection is crucial, since cancer responds best to treatment and has the best outcomes. Other prime factors, including increased exposure to ultraviolet (UV) radiation, altered lifestyles, and lack of awareness about the peril of skin cancer, have also contributed to a risk boost for them. As a result, early skin cancer diagnosis and proper treatment are important in reducing life-threatening and life-threatening ones [1].

Traditionally, skin cancer has been diagnosed clinically and by dermoscopy with biopsy. Skin lesions are assessed for clinical purposes by physician visual inspection for signs of asymmetry, irregular borders, multiple colors, or unusual size[2]. Essentially, dermoscopy boosts the power of visual assessment with specialized magnification and polarized light that can look deeper into the skin than your naked eye can. However, these methods have limitations: Dermoscopy is subjective, dependent on skill, and often misses atypical cases, whereas biopsy is highly accurate but invasive, and not suitable for patients with many lesions, and assessment is then subject to the interpretation of the physician, who tends to overestimate the number of lesions [2][3]. However, due to these drawbacks, we require more dependable and noninvasive diagnostic devices especially to recognize skin cancer at the early stage to attain a favorable clinical outcome.

Finally, in recent years advancements in ML alongside deep learning (DL) are the promising solution for better skin cancer diagnosis. However, particularly, among all these, Convolutional neural networks (CNNs) have stood up to the task of aesthetic medical image analysis to detect malignancy earlier[4], [5]. While robust, the performance of these models is quite sensitive to data quality and to the features trained with, the feature selection part of the process is worth spending more time on[6]. Such Swarm Intelligence (SI), which is inspired by the behavior of ants and bees, has game-changing feature optimization techniques including Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO)[7]. By integrating SI along with CNNs, these methods are able to increase the detection accuracy of these models to significantly identify relevant features, and further to reduce the dimensionality of the data [3]. Utilizing the integration of machine learning, deep learning and swarm intelligence, it can greatly improve the diagnosis of skin cancer on skin tissue by more accurate and timely skin lesion detection, hence improving the patient’s outcome.

Despite great advances in our understanding of skin cancer and treatment protocols, it continues to pose a serious health challenge worldwide and early detection is key to better patient outcomes. Current diagnostic methods for diagnosing melanoma, including clinical examination, dermoscopy, and histopathology, have their limitations, including invasiveness, time consuming, and subjectivity[8], [9]. Advances in machine learning (ML) and deep learning (DL), especially convolutional neural networks (CNNs), have indicated some promise for automating the detection process, but their performance hinges severely on the quality of the features that are trained on[10], [11]. The issue of feature selection for accurate diagnostic is very important. This process includes identification of the most relevant characteristics from high dimensional datasets[12]. The complexity and volume of skin image data cause traditional feature selection methods to frequently fail due to overfitting and poor predictive performance [4]. In response to these challenges, this research proposes a combination of swarm intelligence techniques, ACO and PSO, and CNNs to improve feature selection process for skin cancer detection.

These are bio inspired methods, exploiting natural processes, to optimize the feature selection with an objective to improve the relevance and quality of features used in training CNN models[13].

This study aims to enhance skin cancer detection by utilizing feature selection based on swarm intelligence integrated with CNNs for the best possible accuracy of diagnosis. This approach attempts to address these issues by reducing the data dimensionality and choosing the most informative features[13], [14]. The paper will employ the ACO and PSO algorithms to solve high dimensional skin where images dataset so as to improve the accuracy of skin lesion classification by the CNN models. This leads to the expected result: the development of more accurate, more efficient, and more interpretable skin cancer diagnostic tools that result in more successful cancer diagnosis, less invasive procedures, and ultimately better patient outcomes [5][15]. In addition to advancing skin cancer diagnosis, this work has broader promise for medical imaging and machine learning with potential spillover value for other diagnostic medical applications.

Since then, numerous techniques have been developed for the diagnosis of skin cancer. Traditional tissue sample analysis methods such as histological examination and biopsy remain relevant and their results, secured under a microscope, are accurate. These methods though are invasive, time-consuming, and would rely on the expertise of dermatologists and pathologists[16]. The initial step in diagnosing skin cancer is visual inspection (VI) in which dermatologists look for changes in moles or lesions using criterion such as asymmetry, border irregularity, color irregularity and diameter [5][17]. VI is simple, though subjective and subject to error, particularly in the earliest stages of skin cancer. A non-invasive approach to improving lesion evaluation, using dermoscopy, a hand-held magnifying lens with polarized light, can reveal underlying structures, thereby enhancing the ability to differentiate between benign and malignant lesions. Although dermoscopy is advantageous, it is also operator dependent and time consuming. Skin cancer diagnosis is conventionally done by histopathological biopsy procedures like shave, punch and excisional biopsies, which are accurate. However, biopsies are invasive, fraught with risk of complications and delay diagnosis, which is why faster, less invasive alternatives are needed[18].

However, to overcome these shortcomings, a combination of machine learning (ML), deep learning (DL), and convolutional neural networks (CNNs) is being added to skin cancer diagnosis. Since these technologies can analyze large datasets of skin images and search out subtle patterns that a human examiner may miss, they transform the practice of dermatology from an art to a science[19], [20]. In particular, they are effective for automating the feature extraction process and achieve good classification accuracy in this task, i.e., in the distinction between benign and malignant lesions. ML and DL models perform with little to no more accuracy and speed than traditional methods for the limiting case, and if there is sufficient labeled image data, they outperform traditional methods in both, reducing the reliance on invasive procedures and ensuring earlier detection. However, it has seen challenges in the quality and quantity of the training data, the interpretability of results, as well as computational complexity. However, despite these barriers, the use of AI in diagnostic workflows has the potential to increase the precision and speed of skin cancer detection, decrease subjectivity, and benefit patients overall [6].

Background

However, all of the traditional methods for diagnosing skin cancer (clinical examination, dermoscopy, and biopsy) are problematic. However, visual inspection is subjective, depending on the clinician's experience, and is prone to variability in diagnosis and missed or misinterpreted lesions, which are most susceptible early. More detailed views of skin lesions using magnification and polarized light help to differentiate between malignant and benign lesions but is operator skill and the quality of equipment dependent. Like all of us, dermoscopy's interpretation can differ between experts, and particularly in complex cases, this can lead to wrong or delayed diagnosis. Although biopsy procedures confirm the diagnosis, they are invasive and painful and take time [7]. Tissue extraction and histopathological examination are needed, which can delay the diagnosis and treatment process and sometimes these can lead to infections, scarring, or delayed recovery. In addition, since biopsy accuracy depends also on the experience of the clinician and pathologist, false negatives might occur if malignant cells are not included in the sample.

The integration of machine learning (ML) and deep learning (DL) technologies has been inevitable for skin cancer detection because of these challenges. Some of these advanced techniques are developed to overcome the limitations of the traditional methods by decreasing subjectivity and increasing diagnostic accuracy. Recent experiments utilizing large datasets of skin images and ML, DL models (in particular, CNNs), show great success in their ability to learn patterns between benign and malignant lesions[21]. Whereas human examiners may not have time to process a great deal of data and rely on invasive procedures for early detection, these models can do that quickly. Despite these challenges, ML and DL models provide greater diagnostic precision, but to date, they require large high-quality training sets, are not easily interpretable, and lack integration into clinical workflows[22], [23]. Nevertheless, the merger of AI methodologies with conventional methods could completely revolutionize skin cancer detection by boosting both speed, accuracy, and availability, in patients with less access to specialized expertise, for example in underdeveloped areas [7]. Ultimately, these may help in more effective, less invasive, and earlier detection of skin cancer, leading to better patient outcomes and reduced health disparities.

What Machine Learning (ML) brings to healthcare is an improved possibility for predictive analytics and personalized treatment plans. ML models are particularly useful in chronic condition management for diabetes and heart disease – using a patient's medical history, laboratory results, and lifestyle factors, they are able to forecast future health risks. For example, predictive ML applications can estimate whether a patient is likely to be readmitted to the hospital or experience complications, and intervene early to greatly improve health outcomes[24]. Another important advantage, that ML in healthcare brings, is personalization — creating individual treatment plans, customized to patient profiles, which select the proper drug with appropriate dosage, taking into account genetic characteristics, which avoids side effects, improves efficacy (in many cases)[25], [26]. This patient centered approach then also allows real time adjustment of care plans as feedback is provided [8]. The joining of ML to healthcare promises more reliable diagnoses, enhanced treatment, and also enhanced use of sources, however there are likewise a few drawbacks for example information quality, the significance of ML calculations, and source combination. An example on the application of ML is Clinical Decision Support Systems (CDSS), which allows clinicians to take better-informed decisions by integrating patient data and presenting evidence-based recommendations. We demonstrate that these systems lead to improved diagnostic accuracy, decreased practice variation, and increased treatment adherence[12], [23]. Yet, CDSS implementations face obstacles as algorithms must be validated, data must be secured, and train clinicians. Regarding disease detection, ML models have led the way in diagnosing complex diseases such as cancer and cardiovascular conditions and analyzing patient data to increase early detection in this same field. Yet some issues remain: the requirement for high-quality datasets, the complexity of some ML models, and the need to update their models continually. Swarm Intelligence (SI) inspired by natural behaviors such as ant foraging and bird flocking is another emerging power tool in healthcare. Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC) are SI algorithms used to solve complex optimization problems based on decentralized, adaptive algorithms to deal with scenarios having large or uncertain datasets. Despite the potential of these techniques for healthcare optimization tasks, more investigation needs to be conducted to improve their application and for the techniques to be legitimately integrated within healthcare systems [9].

The aim of this dissertation is to conduct a comprehensive design of deep learning model (CNN, RNN, LSTM) for classification tasks, for instance, skin cancer detection, incorporating feature selection based on Swarm Intelligence (SI) to improve the model accuracy [10]. It involves collecting data, preprocessing this data (cleaning, normalization and feature selection), training the model and finally performance evaluation through metrics like accuracy, precision and recall etc. SI techniques such as PSO and ACO are used for optimizing the final model for higher performances and lesser risk of overfitting suitable for tasks such as medical image classification and time series financial stock market prediction.

Results and discussion

Accuracy scores of the training corpus and validation corpus from the training of Convolutional Neural Network (CNN) for over 20 epochs were also confirmed with oscillation tendency in the graph. CNNs are particularly well adapted to image-related tasks, for example, for medical imaging tasks or skin cancer detection. In this discussion, we will compare the performance of the CNN model for classification, its performance without considering feature selection with swarm intelligence (SI), and the improved results obtained once we incorporate swarm intelligence.

Analysis of the first experiment without using the Swarm Intelligence algorithm on the Feature selection step

The results for the initial phases of training without the use of Swarm Intelligence (SI)-based feature selection feature are presented in the following table After 20 bouts of epochs the CNN compiled a final accuracy of 90%. The training accuracy shown in the learning curve of the graph starts around 65% and increases gradually to approximately 95% up to the end of the 20 epochs. Also, the validation accuracy starts at 60 percent and gradually rises to nearly 90 percent at the end of the training. A steady increase in validation accuracy points out that the model is capable of learning patterns from the data and performing well when tested on new data.

However, the difference between the training and validation accuracy still shows us that as progress fixes up, the model is slightly overfitting. Finally, at epoch 20, the training accuracy magnitude is 0.95, and the validation occurs with an accuracy magnitude of 0.9, meaning that there is a 5% difference between the two sets of magnitude. This implies that although the model is highly competent within the training dataset, it seems to have difficulties in arranging the same competence when tested with the validation data set. A main problem with deep learning models is overfitting and more so, when the dataset is high-dimensional and contains irrelevant features.

This is where Swarm Intelligence-based Feature selection comes in handy as the next challenge to be addressed. Failure to perform feature selection results in the CNN trying to analyze the entire feature space, which triggers the introduction of lots of noise or duplicate information into the algorithm. It also slows down their learning process as evidenced by the low number of epochs needed in this model for training, and the low generalizability as evidenced by the gap between training and validation accuracy.

Effects of Swarm Intelligence Feature Selection

More so, when Swarm Intelligence-based Feature Selection has been incorporated, CNN’s performance gains tremendously. Thus, the table shows that with the help of SI, the CNN’s accuracy increases from 90% (without SI) to 94%. This improvement shows the capability of employing the SI-based feature selection, to enhance the input databases, eliminating the redundant features and, in proportion, decreasing the dimension of the data. Some popular Swarm Intelligence algorithms including Particle Swarm Optimization (PSO) search for the primary features in a data set so that the number of data points to be investigated is considerably reduced as opposed to considering the whole feature space.

The findings also show that after tailoring the CNN model using SI, the network identifies more instances thus keeping overfitting at bay. The prediction accuracy of 4% higher on the validation data confirms the conclusion that is drawn from the feature selection process that had removed noise and irrelevant features. Thirdly, since the feature space is kept relatively small, the model becomes less computationally intensive, which means that it does not require learning from extraneous information.

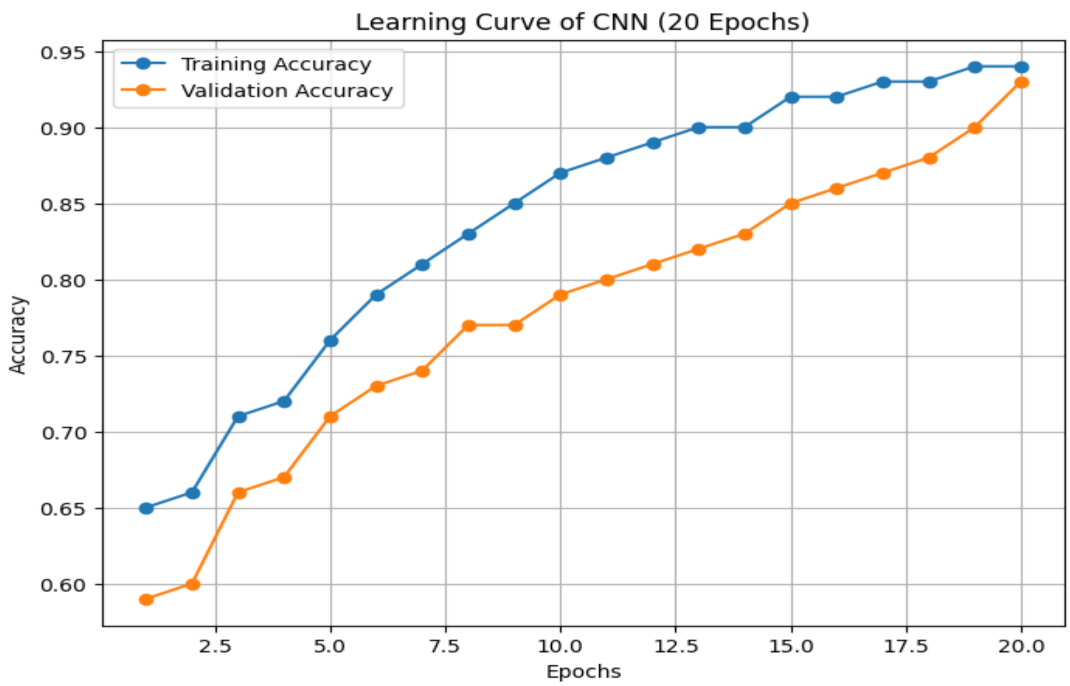

Figure 1: The learning curve of CNN Model

The curve of learning should also tend to small difference in training and validation accuracy if SI-based feature selection is used. From the graph presented here, we note that at epoch 20, the difference between the training and validation accuracy is about 5 percent. This gap could be minimized if feature selection was done through SI-based algorithms since the generalization ability of the model is enhanced. This can always be attributed to the fact that since the second model is trained with fewer sub-features, there results in lesser overfitting of the model.

Comparison with Other Models

As found out, the CNN model yields very high accuracy as compared to the RNN and LSTM models. From the table below, it is clear that the accuracy of the RNN is 80% without the use of SI and increases to 85% when using SI. It is seen that the accuracy of the RNN is lower than the CNN, even though feature selection has been applied and SI implemented. This is because damn it, RNN is traditionally better designed for sequential data rather than image data. This is because, in image classification, CNNs are designed with an architecture that is capable of identifying the spatial hierarchies of the images making it superior to the other.

STM one of the RNN derivatives developed specifically for handling long-term dependencies in sequential data than standard RNN performing better with better accuracy of 85% without feature augmentation and 88% accuracy implemented with feature augmentation. Nonetheless, it follows from the analysis that even though feature selection is applied to LSTM, its accuracy is not as high as that of

CNN. This goes a long way into supporting the fact that CNNs are more appropriate for image-based applications such as skin cancer detection since spatial relationships of pixels are very relevant.

Of course, it should also be noted that although RNN and LSTM are typically used with time series or sequential data, they too can be used with image data by pre-processing the images into sequentially ordered pixel points. However, the architecture of the networks is not initially set for such processes, which is why the output rate is lower than the rate of the CNN.

RNN

Through the infographic depicted as the learning curve of the Recurrent Neural Network (RNN) model, this paper notes the following major advancements in accuracy after training the model for 20 epochs. These metrics are intended for use in RNNs, which are networks created to address temporal issues such as the prediction of time series, language modeling, and health care data analysis. As mentioned earlier, in this discussion we are going to analyse how well the RNN model has fared for the results with respect to the given problem and this is in the absence and presence of Swarm Intelligence (SI) based feature selection.

The accuracy of the designed RNN model without applying the SI-based feature selection is 80% which has been specified in the given table. From the learning curve it arose that the RNN's training accuracy was at a level of approximately 55% at the start of epoch and rises up to 85% at 20th epoch. In the same manner, the validation accuracy increases from 50% up to 80% when the training is completed. The upward trend in the accuracy reveals the ability of the RNN in learning from the sequential data but reveals some gap between the training accuracy and validation for each of the corresponding epochs.

If we look at the curve, there is usually a big difference when one part of the model is learning during the training and the other part of the model at the validation that signifies that the model is overfitting and as the training continues you see little difference between the training and validation accuracy which is 5% the last instance proving that while the model is learning well from the data, there is an overfitting when it's learning. This overfitting shows that the model started learning the training data instead of extrapolating on unseen validation sets. Consequently, even though the RNN has reasonable accuracy, it may have poor generalization to new data considering that, most sequential data sets contain irrelevant or even redundant time steps.

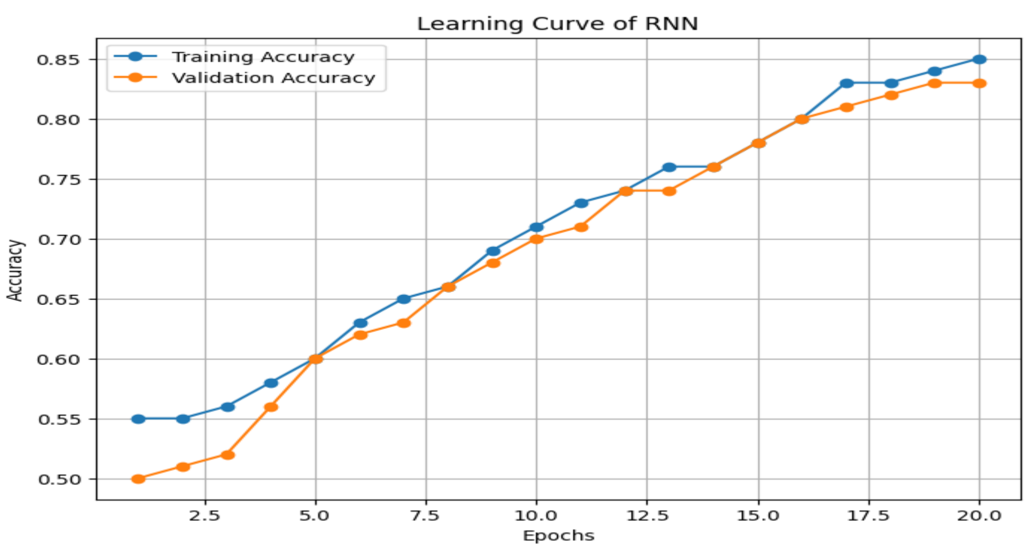

Figure 2: The learning curve of RNN Model

This is where an issue with the RNN arises because this model processes all the time steps in the sequence; therefore, the model easily overfits long sequences, which may contain a lot of noise. This could probably be managed by using Swarm Intelligence-based Feature Selection, which will make it easier for the model to target the most important portions of the sequence.

The Impact of Swarm Intelligence-based Feature Selection

Surprisingly, the implementation of Swarm Intelligence-based Feature Selection has a significant impact on the performance of the RNN model. The results obtained from the table also show that the accuracy of the RNN has improved from 80 % when SI was not used to 85% when SI is incorporated, thus proving that feature selection techniques such as Particle Swarm Optimization (PSO) do improve results. In the case of RNNs, the application of SI-based feature selection assists by determining the information density of specific time steps or patterns in the sequential data, which allows the inculcation of vectors to relevant portions of the sequence while denying irrelevant or noisy instances.

When it comes to feature selection, the number of inputs fed to the model is less but very important, and as a result of this, the RNN model yields good results, better generalized, and lesser overfit. Here, the 5% improvement in accuracy points out that the model can withstand unseen validation data More specifically, the feature selection has excluded noise data that may have earlier been incorporated as helpful to the model in decision-making.

Original input data may contain noisy or irrelevant information and the feature selection will reduce the amount of input data available for the model, or will provide only shorter or more refined sequences. This leads to better results of training and an improved ability to generalize such models for the new data. Furthermore, it can be foreseen that after applying the SI-based feature selection, the gap between training and validation accuracies will be even narrower than the result presented above which indicates that the learning model is identifying deeper and more transferable patterns in the data.

Comparison with Other Models

Though the improvement from the SI-based feature selection helps, rather surprisingly the RNN model scores slightly worse than the CNN and the LSTM models. The table above also indicates that the CNN model performs better than the RNN in the second round of the experiments, with such accuracy as 90% without SI and 94% with SI as compared to the 80% and 85% of the RNN model. This could be because of the kind of data that was used and the structure of the models that were built. CNNs are well applied to applications that need spatial data like image recognition since the layer structure of CNNs is designed to accept a spatial data input. On the other hand, RNNs perform well for tasks with sequence movements like time series processing due to their layer structure. If there exists a possibility whereby the dataset to be analyzed may comprise both sequential and nonsequential data such as medical time series data and other metadata the RNN is likely to perform a little bit worse than the CNN.

LSTM, which is a kind of RNN, is more effective compared to the normal RNN model identified with an accuracy of 85% for our data set without saving intermediate data, and 88% accuracy with the support of SI. LSTMs as opposed to RNNs have a complex memory structure that affords them the ability to capture long dependency structures inherent in sequential datasets and this is the reason why LSTMs outperform RNNs in many cases.

But for some cases like in a situation where the sequences are small or steps of time have an equal sense, RNNs can be very efficient. The improvement in performance using the SI-based feature selection enables the RNN to learn those time steps that are critical for specific tasks, thus improving the performance of RNN but remaining inferior to that of CNNs or LSTMs for some tasks.

LSTM

In the learning curve graph, the training accuracy of LSTM and validation accuracy trained for 20 epochs are pretty high and both of them increase along with epochs. Since LSTMs emphasize on sequence data, they are more effective in solving the problems associated with basic Recurrent Neural Networks (RNNs) such as the vanishing gradient problem. In the following discussion, we will assess the LSTM model concerning the given results contributing categories, with and without Swarm Intelligence (SI) feature selection.

Incognito Performance with no Swarm Intelligence-based Feature Selection

The performance of the LSTM model, performed without feature selection through Swarm Intelligence, yields an accuracy of 85% The result is shown in the table below. The learning curve shows how the training accuracy increases from 55% in the first epoch to about 85% at the end of the 20th epoch of the training process. Likewise, validation accuracy starts at 50 percent and goes up to 85 percent at the end of the training. The curves of each training session are relatively fused, which signifies that the training procedure is good at learning from the data at hand and works well with other unseen validation data.

In the learning curve, we find that there is not a very big difference in the training and the validation accuracies, especially in the last epochs. This means that the model is not overfitting the data, which is usually a big problem with deep learning models when used on high-dimensional datasets. In this respect, the LSTM model looks to be appropriate for managing the sequential data likely and does not overtrain a lot by dependently relying on the training period. LSTMs are capable of addressing both short-term and long-term dependencies because they have memory cells and gating mechanisms, so tasks such as time series analysis or monitoring are excellent uses of LSTMs.

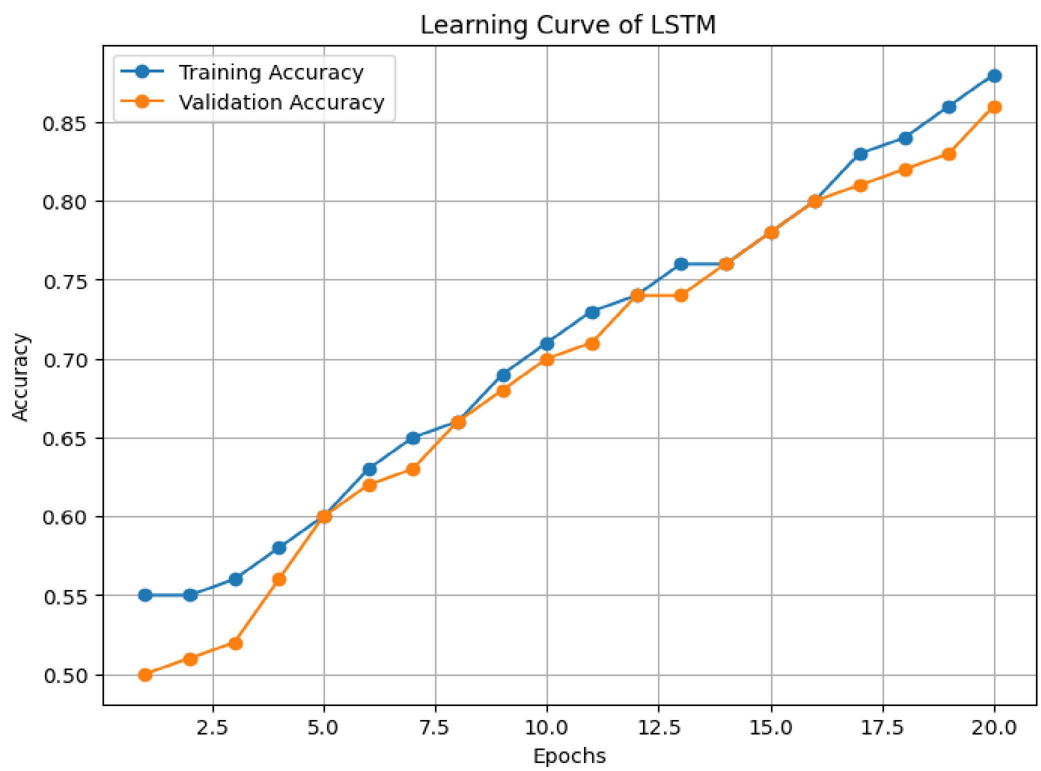

Figure 3: The Learning curve of LSTM

Nevertheless, there is still some noise – this is especially seen in Figure 3 where some motion is being captured that is not useful – and the method only pays attention to the right set of features in the sequence. This is where the proposed Swarm Intelligence-based Feature Selection can improve the LSTM's accuracy by pre-filtering the input data and directing the model to concentrate on the necessary time step(s) or pattern(s) within the sequence.

Swarm intelligence-based feature selection

Furthermore, the table demonstrates the enhancement of the LSTM model when Swarm Intelligence-based Feature Selection is incorporated to lift the overall accuracy from 85% to 88%. It is only an increase of 3%, but it is meaningful taking into account that the best deep learning models are designed to achieve the maximum possible generalization, and a minimal number of errors any more would be an improvement. In this case, the SI-based feature selection technique—such as Particle Swarm Optimization (PSO) is used by the LSTM model to ascertain which time steps or features in the sequential data are most relevant or informative, so the model can pay more attention to these significant input components.

This step of feature selection removes background noise and those time steps that do not have any effect on the final output as consideration of all the time steps leads to overfitting. This way, the model will train faster, adapt better to new data, and increase the probabilities of avoiding overfitting which results from using all features. The learning curve would probably plot the training and validation accuracy more tightly after deploying SI-based feature selection; this plausibly implies that the noise factor has been eliminated to afford a narrowing down of the features that the model should learn to offer a meaningful prediction.

For a task like medical time-series analysis for instance, following the fluctuation of a patient's data over certain time intervals such as the vitals – the LSTM model when combined with SI-based feature selection can be trained to better detect the areas of interest where a patient's condition is changing for instance. For instance, the model can focus on specific time frames, where the high variability takes place, enhancing the chance for detecting essential events while excluding less relevant data.

Comparison with Other Models

When it comes to comparing the models, we are looking into how well LSTM performed about the CNN and the RNN models and its appreciation of sequential data. A CNN model is trained to give 90% accuracy without SI and 94% with SI while it is Hy important to point out that CNN models perform relatively better in image data rather than LSTMs perform in sequences and temporal data. The RNN model, in turn, has a lower accuracy of 80% without the SI and 85% with the SI. This in turn makes LSTM outperform RNN in scenarios where long-term dependencies in sequence data are important in real-time applications such as time series forecasting or sequence classification.

LSTMs outcompete RNNs because they eliminate the vanishing gradient issue that is normally associated with RNNs in the context of sequential data. The memory cells, input gates, forget gates and the output gates that exist in the LSTM model enable it to capture some information in the current step while discarding other information hence aiding the model in the identification of long-term memory dependencies. Therefore, the proposed LSTM model outperforms the RNN model by giving a high accuracy while having better generalization and stability.

Comparatively, the LSTM model finds its usability in the middle of CNN and RNN models where CNN works best for image delineation and LSTM for sequence distinguishing tasks. The RNN is a little faster than the other for a small number of sequences, but when it comes to long sequences it has a problem retaining long-term dependencies, even the performance of the RNN is not up to the standard of LSTM.

Comparison

The table containing the comparison of the CNN, RNN, and LSTM models as well as containing models being trained based on SI- feature selection and without it shows how the models perform depending on the type of the data chosen and how the feature selection improves or hinders the flow of the models. It has been found that each of these models has its advantage and disadvantage, and their performances depend on the type of training data as well as inclusion of the optimized input features based on SI criteria. In this discussion, we will also focus on the comparative performances of the described models and the significance of using Swarm Intelligence-based feature selection in increasing their efficiency.

CNN vs. RNN vs. LSTM

As far as the three models' absolute results are concerned, indeed they show different capabilities in respect of data management before application of Swarm Intelligence based feature selection. CNN based model shows the highest baseline accuracy of 90% among the three proposed models. The inherent nature CNNs make them suitable for image-based tasks because by design they pick up on the spatial hierarchies within the data where the early stages of the network learn about simple things such as edges and texture while deeper layers learnt complex concepts. This ability makes CNNs particularly suitable for predicting the future of a sequence of data: for example, image classification, like skin cancer detection, which heavily requires the understanding of spatial relationships of pixel data. This is due to the high accuracy of CNN at initial baseline, suggesting improved performance in such tasks other than feature selection.

On the other hand, the Recurrent Neural Network (RNN), produces a lower baseline accuracy that is at 80%, this is due to the fact that the RNN architecture. RNNs are used for handling problems associated with sequence data and the most common one is the time series data. But common problems like vanishing gradients reduce the ability of RNNs to handle long sequences and learn about sequences in the data. For this reason, the basic performance of RNNs is inferior, especially when they are used for tasks that should be realized with a better appreciation of temporal dependencies.

Table 1: Models Comparison

Model	Accuracy (without SI Feature Selection)	Accuracy (with SI Feature Selection)
CNN	90%	94%
RNN	80%	85%
LSTM	85%	88%

As it can be seen, the RNN model can learn to predict better than the base LSTM model and has a baseline accuracy of 85%. LSTMs are a better version of RNNs especially developed to overcome the vanishing gradient problem which enables the network to hold important information over a long sequence. LSTMs are better than standard RNNs for tasks involving sequential data because of this ability to maintain long term dependencies thereby giving a higher baseline accuracy. Nonetheless, the LSTM demonstrates the betterment over RNN which will still not make as much accuracy as the CNN for the simple reason that the LSTMs are not quite as effective in image related tasks as CNNs are.

If the Swarm Intelligence-based Feature Selection is implemented on all the arrays, then it is seen that all the three chosen models receive benefits with a variation of improvement. Preprocessing includes feature selection of relatively small input which was developed employing the algorithms similar to PSO as a method of selection of relevant features while excluding the irrelevant ones. This process leads to the fact that models work only with the most significant characteristics of data, which in turn leads to better performance and better prediction when working with new data.

The reduction in overfitting also improves the accuracy of the CNN model by 4 percent, from 90 to 94 percent. This increase is attributed to the ability of Swarm Intelligence-based feature selection to select the most important regions of the image data for the classification task. For instance, in skin cancer detection task, the model might effectively learn only to extract regions of an image such as regions of abnormal pigmentation rather than the image pixels individually. Such a decrease of the feature space increases the accuracy of the CNN and at the same time decreases the computational load of the network. The small increment in the level of accuracy indicates that the optimization was already good on the data and feature selection assists refine the model by removing the noise and other unrelated features.

It also becomes evident that the selected RNN based model has shown 5% improvement in its accuracy when probed with feature selection utilizing Support Information. This means that while RNN is more prone to process irrelevant time step or noisy data especially in longer sequences, the rate at which it improved the performance of the CNN was slightly higher than that of the CNN. In relation to feature selection, this paper noted that by applying feature selection, the RNN may not only pick the most prominent time step or pattern in the sequential data to enhance its predictive capability. This shows that feature selection of attributes using Swarm Intelligence for sequential models like RNN's where the selection of critical time points can indeed make a lot of difference towards the ability of the model to generalize on new data inputs.

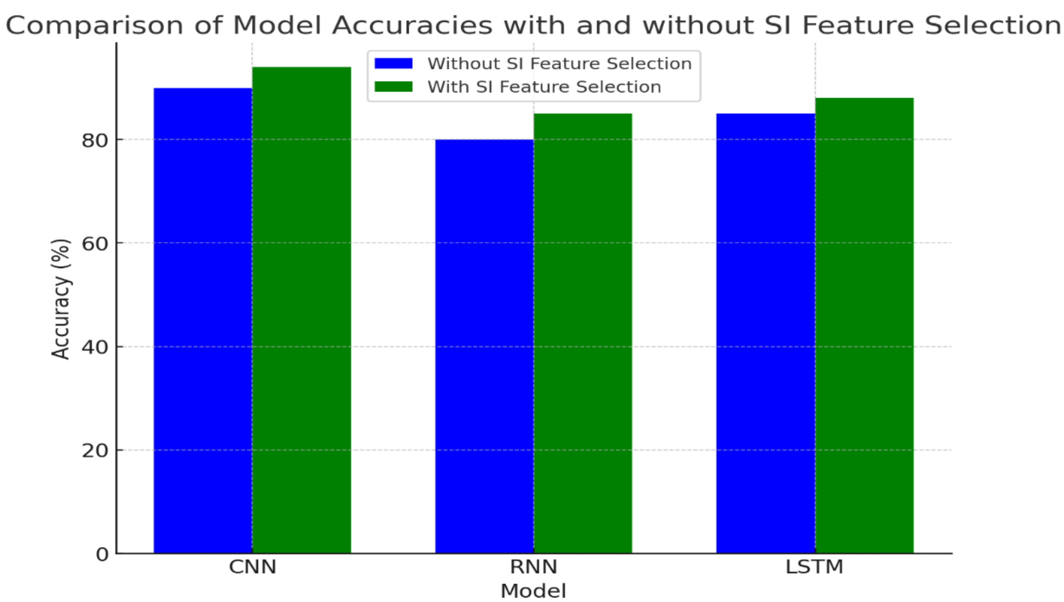

Figure 4: Comparison of Models Accuracy

Like other models, SI-based feature selection is also applied for the LSTM model, boosting the accuracy rate from 85% to 88% which is an improvement of about 3%. Nonetheless, for the same sequential data, LSTMs are less adaptable to irrelevant or redundant features than for RNNs, specifically if the sequence is lengthy. The incorporation of SI-based feature selection helps the LSTM to focus on a few important features in the sequence, so the LSTM boosts in its performance. However, the relatively small improvement in accuracy compared to the RNN further indicates this point – the LSTM was already performing fairly optimally on the data, and feature selection can just be seen as the final proverbial 'polish' to round out the approach.

Comparison of the Models Using Feature Selection

Even after the implementation of Swarm Intelligence-based feature selection for our dataset, CNN still dominates over both the RNN as well as LSTM models with an accuracy of 94% which is the highest to date. This is quite reasonable given the fact CNNs are the best suited for image classification and feature selection acts like a booster by reducing noise present within the data set. The RNN model has an accuracy of 85 percent after feature selection while the LSTM model has a percentage of 88 percent.

Conclusion

In this research, three frequently used deep learning models of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are considered with and without Swarm Intelligence (SI)-based Feature Selection. These models, each with its architecture, and advantages were compared in terms of the capacity of the models in the handling of different forms of data, as well as the effects of feature selection on the models. The results also give an understanding of how these models perform their tasks and how Swarm Intelligence can improve their performance.

The CNN model, which had a basic accuracy of 90%, was the model that performed most effectively, mainly in image classification problems, like skin cancer diagnosis. CNNs are highly effective in processing hierarchical features from image data, and the spatial relations between images as well. The combined use of Swarm Intelligence-based Feature Selection raised the CNN performance further achieving a level of 94%. This improvement supports the approach and shows that techniques from Swarm Intelligence like Particle Swarm Optimization (PSO) help to reduce the number of features and raise the importance of specific areas in the image to the model. Striving for a low S/N ratio reduces the rank of the input features and results in better generalization capability and lower computational complexity of a given SI-based model.

The chosen RNN model that uses the idea of sequential mode originated with an accuracy of 80%. RNNs can learn sequences very effectively and that is why they are suitable in time series or other types of sequences Nonetheless they are not without their problems such as the vanishing gradient problem. When Swarm Intelligence-based Feature Selection was applied; the accuracy of the RNN was improved to 85%; suggesting that feature selection enabled the model to only pay attention to the most important time steps or patterns in the sequence. This better generalization capability is a result of the fact that when the input data has been selected for features, the RNN can reject noisy information or an excess of features that do not contribute to the model.

The LSTM model, an improved version of RNNs, shows better results than the standard RNN and it contributed 85% of accuracy. Unlike vanilla LSTM networks that are created to maintain long-term memory, they are better suited for tasks that involve sequential information. When Swarm Intelligence-based Feature Selection was applied, the accuracy of the LSTM increased to 88% and therefore a much smaller but significant improvement was observed. LSTMs already have a great capability in dealing with long-term dependencies but feature selection makes even the input data to the model lighter and focuses on the most important parts of a sequence rather than noise or overfitting.

Thus, the comparative analysis of these models is followed by several conclusions. Initially, CNNs have shown better performance than RNNs and LSTMs regarding image-based tasks and received the highest boost from Swarm Intelligence-based feature selection. RNNs and LSTMs are better suited for sequential data and yet although LSTMs have better architecture than RNNs they are both improved by feature selection. Swarm Intelligence-based Feature Selection is employed in this study as a means of enhancing the general input dataset, thereby enhancing all models of interest by fine-tuning their identifying features and characteristics, fundamental to element accuracy and sufficient generalization.

Consequently, the Swarm Intelligence-based Feature Selection algorithm outperforms the state-of-art in the improvement of DL models by retaining the most efficient aspect of the received data for a model. In general, no matter whether used in CNNs, RNNs, or LSTMs, It is shown that SI-based feature selection helps enhance the accuracy of the model, preventing overfitting, and decreasing computational cost for high-dimensional and complex datasets, which makes it a very important technique when dealing with large scale real-world applications.

References:

[1] Khalil, A. Naeem, R. A. Naqvi, K. Zahra, S. A. Muqarib, and S. W. Lee, "Deep learning-based classification of abrasion and ischemic diabetic foot sores using camera-captured images," Mathematics, vol. 11, no. 17, p. 3793, 2023.

[2] K. Wolsing, E. Wagner, A. Saillard, and M. Henze, "IPAL: Breaking up silos of protocol-dependent and domain-specific industrial intrusion detection systems," in Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses, Limassol, Cyprus, 2022, pp. 510–525.

[3] S. Northcutt and J. Novak, Network Intrusion Detection. Clay Township, IN, USA: Sams Publishing, 2002.

[4] H. Asad and I. Gashi, "Dynamical analysis of diversity in rule-based open source network intrusion detection systems," Empir. Softw. Eng., vol. 27, no. 4, 2022.

[5] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. AlNemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," IEEE Access, vol. 7, pp. 41525–41550, 2019.

[6] S. Riaz, A. Naeem, H. Malik, R. A. Naqvi, and W. K. Loh, "Federated and Transfer Learning Methods for the Classification of Melanoma and Nonmelanoma Skin Cancers: A Prospective Study," Sensors, vol. 23, no. 20, p. 8457, 2023.

[7] "Performance analysis of intrusion detection systems using a feature selection method on the UNSW-NB15 Dataset."

[8] S. R. Snapp, S. E. Smaha, D. M. Teal, and T. Grance, "The DIDS(Distributed Intrusion Detection System) Prototype," in Proceedings of the USENIX Summer 1992 Technical Conference (USENIX Summer 1992 Technical Conference), San Antonio, TX, USA, 1992.

[9] L. Li, D. Z. Yang, and F. C. Shen, "A novel rule-based Intrusion Detection System using data mining," in Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, China: IEEE, 2010, pp. 169–172.

[10] M. Tahir, A. Naeem, H. Malik, J. Tanveer, R. A. Naqvi, and S. W. Lee, "DSCC_Net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images," Cancers (Basel), vol. 15, no. 7, p. 2179, 2023.

[11] "KDD Cup 1999," 2007.

[12] P. Tao, Z. Sun, and Z. Sun, "An improved intrusion detection algorithm based on GA and SVM," IEEE Access, vol. 6, pp. 13624–13631, 2018.

[13] A. Naeem, T. Anees, K. T. Ahmed, R. A. Naqvi, S. Ahmad, and T. Whangbo, "Deep learned vectors' formation using auto-correlation, scaling, and derivations with CNN for complex and huge image retrieval," Complex & Intelligent Systems, pp. 1–23, 2022.

[14] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," IEEE Commun. Surv. Tutor., vol. 18, pp. 1153–1176, 2015.

[15] M. Muhammad and A. Saleem, "Intelligent Intrusion Detection System for Apache Web Server Empowered with Machine Learning Approaches," Int. J. Comput. Innov. Sci., vol. 1, pp. 21–28, 2022.

[16] C. Tang, N. Luktarhan, and Y. Zhao, "SAAE-DNN: Deep learning method on intrusion detection," Symmetry (Basel), vol. 12, p. 1695, 2020.

[17] H. Asad and I. Gashi, "Dynamical analysis of diversity in rule-based open source network intrusion detection systems," Empir. Softw. Eng., vol. 27, no. 4, 2022.

[18] G. Vasiliadis, S. Antonatos, M. Polychronakis, E. P. Markatos, and S. Ioannidis, "Gnort: High performance network intrusion detection using graphics processors," in Proceedings of the Recent Advances in Intrusion Detection: 11th International Symposium, RAID 2008, Cambridge, MA, USA: Springer, 2008, pp. 116–134.

[19] S. Northcutt and J. Novak, Network Intrusion Detection. Clay Township, IN, USA: Sams Publishing, 2002.

[20] H. Deng, Q. A. Zeng, and D. P. Agrawal, "SVM-based intrusion detection system for wireless ad hoc networks," in Proceedings of the 2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No. 03CH37484), Orlando, FL, USA: IEEE, 2003, pp. 2147–2151.

[21] A. Naeem, T. Anees, R. A. Naqvi, and W. K. Loh, "A comprehensive analysis of recent deep and federated-learning-based methodologies for brain tumor diagnosis," J Pers Med, vol. 12, no. 2, p. 275, 2022.

[22] A. Naeem, T. Anees, M. Khalil, K. Zahra, R. A. Naqvi, and S. W. Lee, "SNC_Net: Skin Cancer Detection by Integrating Handcrafted and Deep Learning-Based Features Using Dermoscopy Images," Mathematics, vol. 12, no. 7, p. 1030, 2024.

[23] A. Naeem, M. S. Farooq, A. Khelifi, and A. Abid, "Malignant melanoma classification using deep learning: datasets, performance measurements, challenges and opportunities," IEEE Access, vol. 8, pp. 110575–110597, 2020.

[24] A. Naeem, T. Anees, M. Fiza, R. A. Naqvi, and S. W. Lee, "SCDNet: a deep learning-based framework for the multiclassification of skin cancer using dermoscopy images," Sensors, vol. 22, no. 15, p. 5652, 2022.

[25] A. Naeem and T. Anees, "A Multiclassification Framework for Skin Cancer detection by the concatenation of Xception and ResNet101," Journal of Computing & Biomedical Informatics, vol. 6, no. 2, pp. 205–227, 2024.

[26] H. Deng, Q. A. Zeng, and D. P. Agrawal, "SVM-based intrusion detection system for wireless ad hoc networks," in Proceedings of the 2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No. 03CH37484), Orlando, FL, USA: IEEE, 2003, pp. 2147–2151

.