Regional Rainfall Prediction Using Support Vector Machine Classification of Large-Scale Precipitation Maps

Eslam Hussein
Department of Computer Science
University of the Western Cape
Cape Town, South Africa
eslamhuss34@gmail.com

Mehrdad Ghaziasgar
Department of Computer Science
University of the Western Cape
Cape Town, South Africa
mghaziasgar@uwc.ac.za

Christopher Thron
Department of Science and Mathematics
University-Central Texas
Killeen, Texas, USA
thron@tamuct.edu

Abstract—**Rainfall prediction helps planners anticipate potential social and economic impacts produced by too much or too little rain. This research investigates a class-based approach to rainfall prediction from 1-30 days in advance. The study made regional predictions based on sequences of daily rainfall maps of the continental US, with rainfall quantized at 3 levels: light or no rain; moderate; and heavy rain. Three regions were selected, corresponding to three squares from a 5×5 grid covering the map area. Rainfall predictions up to 30 days ahead for these three regions were based on a support vector machine (SVM) applied to consecutive sequences of prior daily rainfall map images. The results show that predictions for corner squares in the grid were less accurate than predictions obtained by a simple untrained classifier. However, SVM predictions for a central region outperformed the other two regions, as well as the untrained classifier. We conclude that there is some evidence that SVMs applied to large-scale precipitation maps can under some conditions give useful information for predicting regional rainfall, but care must be taken to avoid pitfalls.**

Index Terms—**a comparison study, a sequence of images, SVMs**

I. INTRODUCTION

A. Role of rainfall maps in water resource management

Rainfall maps provide essential information about intensity, temporal, and spatial location which are essential in water resource management. Historical rainfall maps data can help different management sectors such as agriculture to make informed decisions about water supply management strategies to better utilize the occurrence of precipitation events [1]. Historical data can be most effectively utilized by developing prediction models such as machine learning to capture historical rainfall patterns.

In previous literature, prediction models based on rainfall maps may be grouped into two main categories. The first type involves applying deep learning to a sequence of images as an input to predict future frames. Usually, the images used for this type of prediction are separated by relatively small time intervals e.g 6-10 minutes [2]–[8]. The second type consists of single output regression or classification-based models. These models predict rainfall on an hourly [9]–[11], daily [12], [13],

or monthly [14] basis using prior rainfall maps. Regression-based models may use a single frame [14] or batch of frames [10] as an input to give a numerical rainfall prediction. In contrast, classification based models categorize local rainfall into two or more discrete classes and predict the classes of future precipitation events based on a single frame or a batch of frames. [12], [13], [15]. Both regression and classification models can be used to predict entire images one pixel at a time [11], [16].

For purposes of comparison, we describe the work of Boonyuen in [12] and [13]. In [12] the authors used a single image to produce a binary classification (rain/no-rain) for three days ahead in Thailand. Using the inception-V3 based CNN model the authors had up to 54.84% classification accuracy for three days ahead prediction. The study also concluded that including neighboring countries in the images increases the efficiency of the model compared to cropping the image to focus only on Thailand. In [13] the authors developed an inception-V3 model to classify predicted rainfall into four categories (No-rain, light-rain, moderate-rain, heavy-rain). Both batches of satellite images and single images were used as input. The study demonstrated that using batches of images as input makes the model more robust at classifying upcoming rainfall. The trained model was able to predict one, two, three days ahead with an accuracy of 70.58%. Having the same accuracy up to three days ahead is an issue of concern, as we suspect that the trained model has a bias towards the majority class (no-rain), making the accuracy to be constant. Measuring the efficiency of models using classification accuracy on imbalanced data is not ideal, because the results obtained may reflect the relative frequencies of the classes more than the actual effectiveness of the model. When imbalanced classes are involved, the f1-score can be a better measure of the method's effectiveness [17], as it takes the weighted average of precision and recall and is less influenced by class imbalance.

In The literature, various image sizes and sequence lengths are used in different prediction models. It appears that these parameters are usually chosen arbitrarily, or determined by
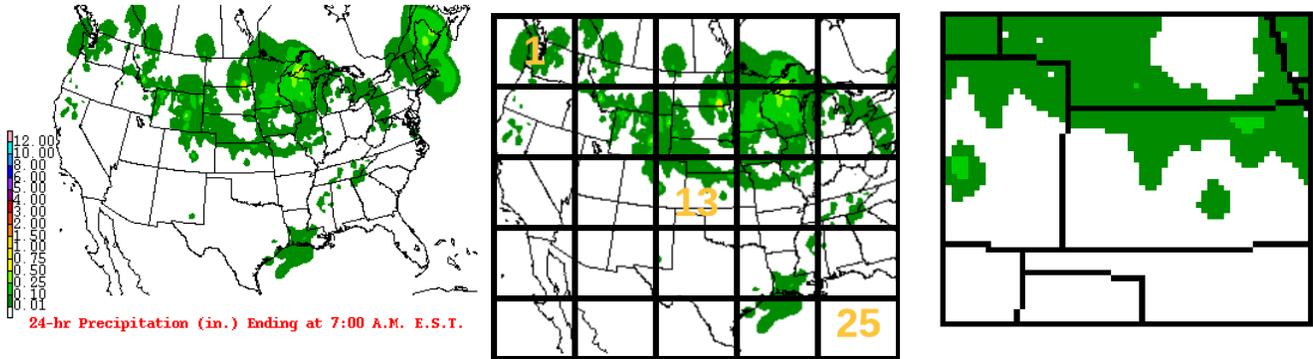
Fig. 1. Satellite precipitation images: (left) An image from the NCEP dataset; (centre) 5 grid overlay; (right) tile 13.

trail and error. For example, the authors in [10] used up to 60 images to predict rainfall on an hourly base. the authors in [3], [7], [8], [18]–[20], used 4, 4, 5, 10, 10, 20, .20 images respectively. As to image size, the authors in [10] used several image sizes between 101×101 and 10×10, and compared the performance.

Even though most previous studies made use of deep learning related techniques to capture rainfall patterns on the historical images like Convolutional LSTM in [2], in some cases the deep learning approach has significant drawbacks. In particular, as it often overfits when the training set is relatively small [21], [22]. In addition, those models have many hyperparameters that need to be optimized.

These difficulties can be avoided by using a support vector machine (SVM) approach instead of deep learning. SVM is a powerful machine learning technique where it is often used in classification and regression problems [23]. SVM is a classifier that generates a hyperplane to classify data instances [24], [25], where optimal hyperplanes are determined by constructing the largest margin of separation between the different instances [26]. In contrast to deep learning, SVMs are suited to be trained on small and medium-sized complex datasets [23], [27]. SVM also has the capability to perform structural risk minimization (SRM), which enables SVMs to avoid overfitting by minimizing the bound on the generalization error [28].

Using SVMs can also avoid the need for extensive hyperparameter tuning. For example, the authors in [29] considered the use of linear kernel (SVMs) in case of having the number of features exceeds the size of the dataset. Linear kernels use only a single parameter, the regularization parameter C, that determines the trade-off between minimizing the training error and the model complexity [30].

Several recent papers in the literature use SVMs on rainfall prediction for different classifications and regression problems. The authors in [31] investigated the use of SVMs as well as other techniques to classify rainfall on a very small training set 10% (2245), where the output was a binary classification rain/no-rain daily. Another binary classification problem (rain/no-rain) was studied in [25], which investigated the use

of SVMs on weather stationary data to classify rainfall for the next five minutes. The data were highly imbalanced due to the rare occasion of rain, which made the researchers perform down-sampling on the dataset. As for regression problems, The authors in [32] used SVM to predict daily and the accumulated rainfall on 42 different cities from Europe and the US. The authors in [11] investigated the use of SVMs with hourly radar-derived rainfall to predict precipitation during typhoons. Another study linked the observations from satellite imagery data to predict rainfall up to 6-hours [15].

B. Scope of this research

This research aims to investigate precipitation forecasting on a dataset from the National Center for Environmental Prediction (NCEP) using SVMs. Our investigation has three aspects: i) Determine the effect of image sequence input length on class prediction accuracy, ii) Assess the effect of image size on class prediction accuracy, and iii) Compare the accuracy of rainfall class predictions for three selected squares (tiles) from a 5×5 grid covering the map area, for up to 30 days ahead.

This paper is organized as follows. Section II presents the methodology, including a discussion of the datasets and their preparation as well as the SVM specification and training. Section III presents the results in tabular and graphical form and provides analysis and discussion. Section IV summarizes our conclusions.

II. METHODOLOGY

Following the flow chart in Figure 2, we start by discussing the data set, followed by the pre-processing of the images and the preparation of the data then explaining one of the models for this prediction.

A. Data Set

1) 24-Hour-Precipitation-Forecasting: The data used for this study are radar images taken daily at 7 a.m. Eastern Standard Time, from Jan 2012 to Oct 2019, with a total of 2,835 images. The data comes from the NCEP, with a size of 400×320 pixels which represents the United States. Each image contains 16 different rainfall intensity level, Figure 1 (Left) shows a full image of the used dataset.
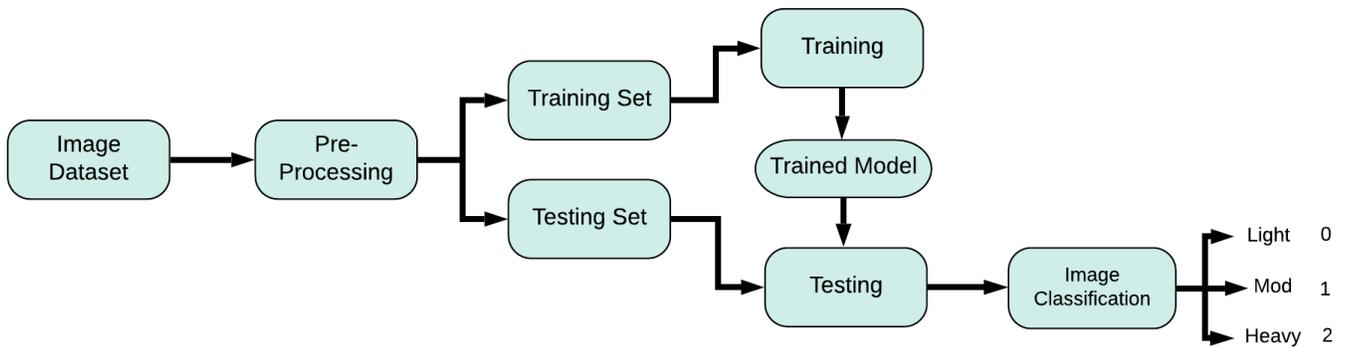
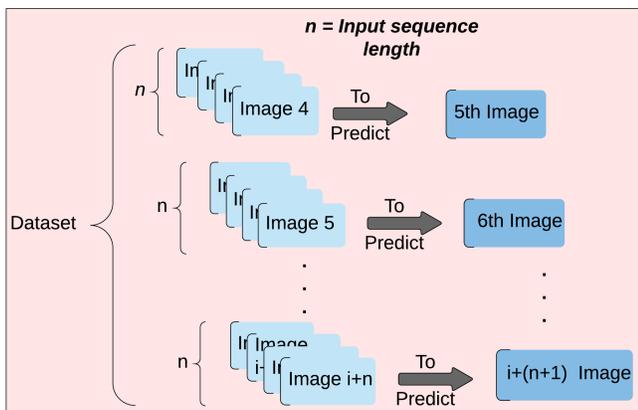Fig. 2. Flow chart showing the implementation process.

Fig. 3. Overview of dataset preparation.

B. Pre-Processing

The input data set went through a pre-processing transformation as shown in Figure 2. The first transformation included size reduction, as we reduced the size of the images from 400×320 to 172×123 (image size a), and 87×61 (image size b). For both reduced data sets, we cropped the images to remove some of the irrelevant information that exists in our data set like the color bar on the left of Figure 1 (left map). In this study, we did not consider the full image size due to memory and time consideration. After resizing we transformed the images into one channel by performing a grayscale transformation on the images to reduce the complexity of the model.

Figure 3, shows how we prepared the input dataset in a sliding window fashion to predict the next image in a sequence. Our use of sliding windows resembles the approach in several previous references [8], [20], [32]–[34]. The figure shows the case where $n = 4$, which stands for the size of the window (sequence length). In general, the number of feature (pixels) of one sample of the data set can be determined by $n \times w \times h$, where n is the size of the window, w is the

width of the image, and h is the height of the image, which depends on the size of the images. For optimization, the study tested different size window sizes n, where $n \in \{2, 4, 6, 8\}$. We divided that data into 90% training and 10% testing for the whole experiment. Moreover, we divided the US map into a 5×5 grid squares (tiles), as We trained on specific grid tiles which are 1, 13, 25.

Initially, we divided the 16 classes on the color bar by Figure 1 left image equally into three classes light, moderate, heavy. For each image, tiles were classified according to the highest level observed within the tile: for example, if the tile had one or more pixels showing heavy rainfall, the entire tile was classified as heavy rain for that image. However, this equal division produced highly unbalanced data, due to the rare occasion of very heavy rainfall. Consequently, our classification accuracy on the testing test was constant for predictions up to 30 days ahead, which is similar to what was observed in [13] as discussed in the Introduction. To circumvent this problem, we made an unequal division between the classes by designating the lowest three classes as no/light rain, the next three classes as moderate rain, and all remaining classes as heavy rain. This improved the balance between the three classes: for the three tiles, we observed the following frequencies (no/light rain, moderate rain, heavy rain): (25%, 49%, 26%) for tile 1, (38%, 30%, 32%) for tile 13, and (36%, 37%, 27%) for tile 25.

The three tiles show quite different seasonal behavior, as shown in Figure 4. For Tile 13 (central), there is a clear distinction between the light rain and heavy rain class frequencies between summer months (May-Aug) and winter months (Oct-Mar). For Tile 25 (southeast), the light/no rain class shows strong seasonality, while the other two classes less so. For Tile 1 (northwest) the seasonality for all classes is less distinct.

To summarise, the input to our model is a set of full images, with different images sizes and windows (sequence length), while the prediction of the rainfall intensity happens on three specific tiles corresponding to three local regions within the U.S.

TABLE I

TILE 1 (F1-SCORE ACCURACY ON THE TESTING SET). WITH k DAYS AHEAD (DA), WITH DIFFERENT n INPUT IMAGES AND SIZES.

Image Scale	Input Images	F1-score of Days Ahead (DA) (%)									
		1DA	2DA	3DA	4DA	5DA	6DA	7DA	14DA	30DA	Mean
Img(size *a*)	2	**60**	31	27	30	26	23	**34**	22	22	30.55
	4	52	37	**40**	26	24	32	22	22	22	30.77
	6	50	37	37	32	31	22	22	22	22	30.66
	8	44	31	30	29	27	25	22	22	22	28
Img(size *b*)	2	55	33	40	25	25	**33**	30	22	22	31.66
	4	57	**43**	30	**35**	**37**	25	22	22	22	**32.55**
	6	51	42	32	32	32	22	22	**29**	**23**	31.66
	8	53	34	36	35	24	25	23	22	22	30.44
Mean		52.75	36	34	30.5	28.25	25.875	24.625	22.875	22.125	

TABLE II

TILE 13 (F1-SCORE ACCURACY ON THE TESTING SET). WITH k DAYS AHEAD (DA), WITH DIFFERENT n INPUT IMAGES AND SIZES.

Image Scale	Input Images	F1-score of Days Ahead (DA) (%)									
		1DA	2DA	3DA	4DA	5DA	6DA	7DA	14DA	30DA	Mean
Img(size *a*)	2	51	49	**49**	42	42	42	43	38	44	44.44
	4	58	**50**	45	41	44	**44**	45	46	**45**	**46.44**
	6	51	50	42	39	44	44	40	**48**	41	44.33
	8	54	45	46	43	44	44	46	45	44	45.66
Img(size *b*)	2	55	46	48	**46**	41	40	**48**	46	37	45.22
	4	55	47	43	45	43	41	43	43	40	44.44
	6	57	48	46	43	**49**	43	43	42	41	45.77
	8	**60**	49	48	42	44	41	45	42	45	46.22
Mean		55.125	48	45.875	42.625	43.875	42.375	44.125	43.75	42.125	

TABLE III

TILE 25 (F1-SCORE ACCURACY ON THE TESTING SET). WITH k DAYS AHEAD (DA), WITH DIFFERENT n INPUT IMAGES AND SIZES.

Image Scale	Input Images	F1-score of Days Ahead (DA) (%)									
		1DA	2DA	3DA	4DA	5DA	6DA	7DA	14DA	30DA	Mean
Img(size *a*)	2	41	40	40	37	40	46	**49**	44	39	41
	4	45	35	46	36	43	44	36	40	**49**	41.55
	6	46	48	39	46	39	41	39	42	49	43.22
	8	53	37	41	46	42	41	42	39	42	42.55
Img(size *b*)	2	48	41	39	37	41	42	41	**45**	34	40.88
	4	42	38	42	**48**	41	**47**	41	42	42	42.55
	6	**56**	**50**	**49**	44	**46**	39	41	41	41	**45.22**
	8	49	42	42	46	42	43	41	43	42	43.33
Mean		47.5	41.375	42.25	42.5	41.75	42.875	40.375	42	42.25	

C. Classification

In this study, we used linear kernel SVMs, based on [29] which advised using a linear kernel when the number of features exceeds the size of the dataset. These were trained on the training set, which comprised 90% of the data. Training was accomplished using the *sklearn.svm.SVC* class in scikit-learn (www.scikit-learn.org). Since we are working with a linear kernel, we had only one parameter to optimize which is the regularization parameter C. We tried values $C = 2^k$ for $k \in -15, ...6$ with 10-fold cross-validation on the training set, and we chose the best C value separately for each input configuration and each day ahead prediction, for each of the three tiles. (It was observed that in most cases the value of C thus obtained was in the 2^{-13}–2^{-10} range.) The models obtained were then applied to the testing set, and confusion matrices were obtained which were used to compute macro f1 scores. The macro f1-score was calculated as the average of the f1 scores of each class, where the per-class f1 score is computed as follows:

$$\text{f1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{1}$$

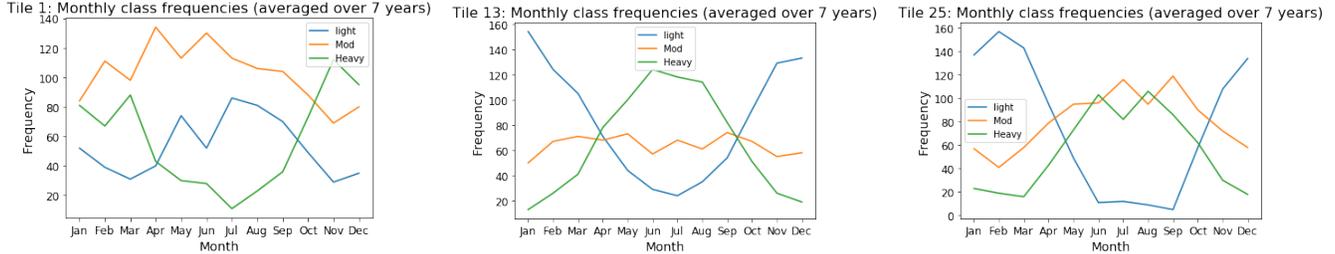| Region	F1-score range of Days Ahead (DA) (%)								
	1DA	2DA	3DA	4DA	5DA	6DA	7DA	14DA	30DA
tile 1	16	12	13	10	13	11	12	7	1
tile 13	9	5	7	7	8	4	8	10	8
tile 25	15	15	10	12	7	8	6	6	15

Fig. 4. Seasonality on a monthly basis for Tile 1 (*left*), Tile 2 (*middle*), and Tile 3 (*right*).

III. RESULTS AND DISCUSSION

Tables I, II,III show the macro f1-score accuracy on the testing set for tile 1 (northwest), tile 13 (central) and tile 25 (southeast) respectively. Each table gives results for different days-ahead predictions for 8 different SVM inputs (four different input sequences and two different image sizes). The bold numbers in each column represent the highest macro f1-score among the 8 SVM inputs for that specific days-ahead prediction for the given tile.

In the following discussion, we will first compare the prediction performance for the different SVM inputs. Then we will compare the prediction performance for the three different geographical regions.

A. Comparison between different SVM inputs

From Tables I, II, and III we do not find that any one input configuration is clearly better than the others. In Table I for example, we find that 5 different input configurations attain the best accuracy for different days-ahead predictions. There is considerable variation within each column of the tables, as well as from column to column for each row. Table IV shows the f1 score range (maximum − minimum) among the 8 predictions for each days-ahead, for the three regions. From the table, it is clear that Tile 13, in general, has the lowest variability in the 1-6 day range, while the variability for Tile 1 reduces to almost 0 after 30 days.

The observed variabilities may be attributed at least partially to the relatively small size of the testing set, which consists of about 300 images. For purposes of comparison, a sequence of 300 Bernoulli trials with success probability $p = 0.5$ will have a standard deviation of ± 3 percentage points. So a 95% confidence interval of ± 2 standard deviations will have a width of 12 percentage points. So in this Bernoulli trial scenario, the probability of getting 8 independent trials within a range of 12 is roughly $0.95^8 = 0.66$.
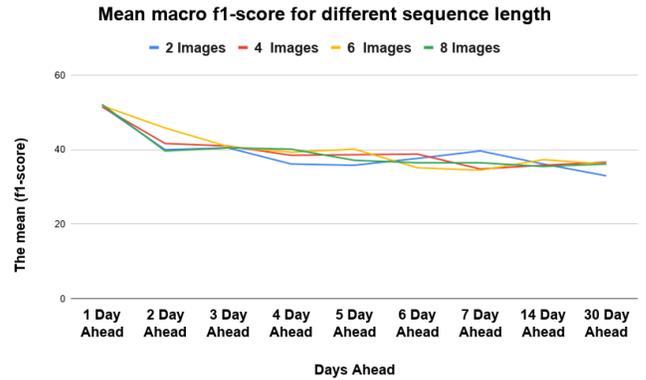

Fig. 5. Mean macro f1-scores as a function of days ahead for different input image sequence length, averaged over all image sizes and tiles.

Figures 5 and 6 isolate the effect of input sequence length and image size, respectively. In Figure 5 the macro f1 scores for all predictions for all tiles for each days-ahead were averaged, and the results plotted as a function of days-ahead. The figure shows that no particular input sequence length is superior to the others. Figure 6 similarly averages macro f1 scores separately for each image size. There appears to be a slight advantage of about 1 percentage point when using the larger image size (172×123) instead of the smaller size (87×61). Both figures show a clear decrease in accuracy as days-ahead increases. in contrast to the constant classification accuracy found in [13].

B. Comparison between regional predictions

Figure 7 shows the macro f1 scores averaged over 8 inputs (4 sequence lengths $\times$ 2 images sizes) for each day ahead, for each image separately. These f1 scores are compared to f1 scores obtained from an untrained predictor that simply
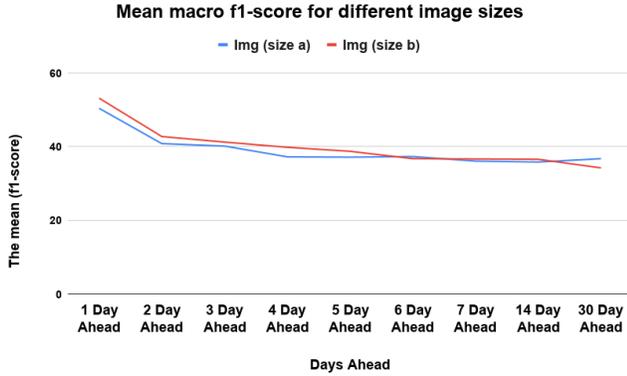
Fig. 6. Mean macro f1 score as a function of days ahead for the two different image sizes, averaged over all input sequence lengths and tiles.

Fig. 7. Per-tile average of f1-scores for 8 SVM inputs as a function of days ahead, compared to untrained benchmark predictor.

uses the class of the final image in the sequence as the predicted class for $1, 2, 3, \ldots 30$ days ahead. From the figure, it is clear that the SVM significantly underperforms the untrained benchmark predictor for Tiles 1 and 25. For Tile 13, the SVM outperforms the untrained predictor by about 1-3 percent.

The SVM performance for Tile 1 is particularly poor, especially for longer-range predictions. Upon closer examination of the confusion matrices produced by the simulation, we found that for predictions longer than 4 days ahead, almost all of the SVM inputs were producing a classifier that always predicted the most frequent class (moderate rain) regardless of the SVM input. An elementary calculation based only on class frequencies (25%, 49%, 26%) shows that a predictor which ignores inputs and always chooses the majority class will have a macro f1-score of 21.5%, which is consistent with our observed result.

The SVM performance for Tile 25 is also worse than the untrained predictor. The confusion matrices for Tile 25 shows that for larger-DA predictions, most predictors simply choose between light and moderate classes, and never predict the heavy class. This result is understandable based on the seasonal behavior of Tile 25, shown in Figure 4 (*right*). During the winter months (Nov.-Dec.), the light class is by far the most frequent. So if the input images show the tile as belonging to the light class, the chances are that the class will remain light for the next 30 days. On the other hand, in the summer months (May-Sep), the light class is virtually nonexistent, and moderate and heavy are about equal. Since moderate is a more frequent class than heavy (37% versus 27%), the classifier favors moderate over heavy. Apparently, the untrained benchmark predictor more closely matches the seasonal pattern, which is why the untrained predictor performs better. We note that for Tile 25, a predictor that always chooses the majority class will have a macro f1 score equal to 18%, so the SVM does represent a large improvement over a majority-class predictor.

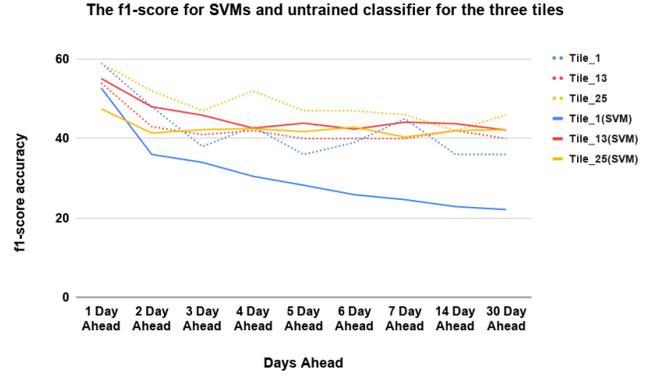Tile 13 showed the best SVM performance among the three tiles and was the only tile where the SVM outperformed the untrained benchmark classifier. We also noted from Table IV that there was a smaller variation in macro f1 scores among the 8 different SVM inputs, for predictions between 1-5 days ahead. It is reasonable that the SVM's in Tile 13 are finding useful features and converging, while the SVM's for the other tiles are not locating truly useful features, so they are overfitting the training set which means that they no longer give consistent accuracy when applied to the testing set. It may be argued that Tile 13 has better input data than the other tiles because the input images contain information from all regions surrounding Tile 13 which is not the case for the other two tiles.

IV. CONCLUSION

Our study points out some of the limitations and potential pitfalls in using SVMs with linear kernels for weather prediction up to 30 days in advance. We have shown that some classifiers used by previous researchers (e.g. [12] and [13]) which seem to show good performance may be largely due to unequal class divisions rather than the classifier itself. We have also shown that unequal classes may cause linear SVMs to converge on majority-class classifiers, or to completely neglect classes of low frequency. Among the three geographical regions predicted, only the central region had an SVM-based classifier that performed better than a simple untrained classifier that used the tile's class in the final image of the input sequence as the prediction. We conjecture that SVMs may be performing better on the central tile because the input sequence contains precipitation information for all surrounding tiles, which is not the case for tiles at the corners of the map (such as Tile 1 and Tile 25).

In this study, we propose to divide the map of the US onto 25 tiles, as for the optimization we used different input configurations. The support vector machine was used to classify the image sequences as no/ light-rain: a) Our study shows that f1-score for tile 13 has a generally better f1-score than tile 1 and tile 25, which goes back to the position of the tile. b) Taking a bigger size scale as an input appears to provide slightly better performance than the smaller scale, but with higher time cost.

c) The f1-score shows a decay while predicting days ahead, but the decay does not appear as prominent with imbalanced tiles e.g tile 1.

We have argued that the variability in the results obtained from day to day and from input method to input method was at least partially due to the insufficient amount of data for training and testing. Much of the variation in observed f1 scores is attributable to the small testing set size. Recent research has shown that using data augmentation to augment the set of training images may improve the efficiency of the trained model [5]. With an augmented training set, a larger portion of the actual data may be used for testing.

Another significant drawback of the SVM classifiers used was that they did not take seasonality into account. This is why a simple untrained classifier that took advantage of seasonality was able to substantially outperform the SVM classifiers on two out of the three regions examined. It is possible that SVMs that take seasonality into account my perform much better: this is a possible area for future research.

In the current research we did not attempt to include additional engineered features, because the number of features used was already very large. For future investigation, a possible approach would be to use PCA to reduce the number of features, and then apply feature engineering. In addition to time stamp, Gaussian Mixture Models may be used to capture spatial means and variances, which could be used as global features.

REFERENCES

[1] R. Harmel, K. King, C. Richardson, and J. Williams, "Long–term precipitation analyses for the central texas blackland prairie," *Transactions of the ASAE*, vol. 46, no. 5, p. 1381, 2003.

[2] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems*, 2015, pp. 802–810.

[3] E. Shi, Q. Li, D. Gu, and Z. Zhao, "A method of weather radar echo extrapolation based on convolutional neural networks," in *International Conference on Multimedia Modeling*. Springer, 2018, pp. 16–28.

[4] X. Shi, Z. Gao, L. Lausen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. Woo, "Deep learning for precipitation nowcasting: A benchmark and a new model," in *Advances in neural information processing systems*, 2017, pp. 5617–5627.

[5] Q.-K. Tran and S.-k. Song, "Multi-channel weather radar echo extrapolation with convolutional recurrent neural networks," *Remote Sensing*, vol. 11, no. 19, p. 2303, 2019.

[6] Y. Wang, M. Long, J. Wang, Z. Gao, and S. Y. Philip, "Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms," in *Advances in Neural Information Processing Systems*, 2017, pp. 879–888.

[7] S. Singh, S. Sarkar, and P. Mitra, "A deep learning based approach with adversarial regularization for doppler weather radar echo prediction," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2017, pp. 5205–5208.

[8] L. Chen, Y. Cao, L. Ma, and J. Zhang, "A deep learning based methodology for precipitation nowcasting with radar," *Earth and Space Science*, p. e2019EA000812, 2020.

[9] S. Kim, S. Hong, M. Joh, and S.-k. Song, "Deeprain: Convlstm network for precipitation prediction using multichannel radar data," *arXiv preprint arXiv:1711.02316*, 2017.

[10] c. zhang, H. Wang, J. Zeng, L. Ma, and L. Guan, "Tiny-rainnet: A deep cnn-bilstm model for short-term rainfall prediction."

[11] P.-S. Yu, T.-C. Yang, S.-Y. Chen, C.-M. Kuo, and H.-W. Tseng, "Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting," *Journal of Hydrology*, vol. 552, pp. 92–104, 2017.

[12] K. Boonyuen, P. Kaewprapha, and P. Srivihok, "Daily rainfall forecast model from satellite image using convolution neural network," in *2018 IEEE International Conference on Information Technology*, 2018, pp. 1–7.

[13] K. Boonyuen, P. Kaewprapha, U. Weesakul, and P. Srivihok, "Convolutional neural network inception-v3: A machine learning approach for leveling short-range rainfall forecast model from satellite image," in *International Conference on Swarm Intelligence*. Springer, 2019, pp. 105–115.

[14] S. Aswin, P. Geetha, and R. Vinayakumar, "Deep learning models for the prediction of rainfall," in *2018 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2018, pp. 0657–0661.

[15] K. Chen, J. Liu, S. Guo, J. Chen, P. Liu, J. Qian, H. Chen, and B. Sun, "Short-term precipitation occurrence prediction for strong convective weather using fy2-g satellite data: a case study of shenzhen, south china," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 41, p. 215, 2016.

[16] A. Mukhopadhyay, B. P. Shukla, D. Mukherjee, and B. Chanda, "A novel neural network based meteorological image prediction from a given sequence of images," in *2011 Second International Conference on Emerging Applications of Information Technology*. IEEE, 2011, pp. 202–205.

[17] L. A. Jeni, J. F. Cohn, and F. De La Torre, "Facing imbalanced data–recommendations for the use of performance metrics," in *2013 Humaine association conference on affective computing and intelligent interaction*. IEEE, 2013, pp. 245–251.

[18] E. Shi, Q. Li, D. Gu, and Z. Zhao, "Convolutional neural networks applied on weather radar echo extrapolation," *DEStech Transactions on Computer Science and Engineering*, no. csae, 2017.

[19] L. Tian, X. Li, Y. Ye, P. Xie, and Y. Li, "A generative adversarial gated recurrent unit model for precipitation nowcasting," *IEEE Geoscience and Remote Sensing Letters*, 2019.

[20] Y. Cao, Q. Li, H. Shan, Z. Huang, L. Chen, L. Ma, and J. Zhang, "Precipitation nowcasting with star-bridge networks," *arXiv preprint arXiv:1907.08069*, 2019.

[21] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in *2015 3rd IAPR Asian conference on pattern recognition (ACPR)*. IEEE, 2015, pp. 730–734.

[22] P. Liu, K.-K. R. Choo, L. Wang, and F. Huang, "Svm or deep learning? a comparative study on remote sensing image classification," *Soft Computing*, vol. 21, no. 23, pp. 7053–7065, 2017.

[23] A. Géron, *Hands-on machine learning with Scikit-Learn and Tensor-Flow: concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.", 2017.

[24] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[25] S. Manandhar, S. Dev, Y. H. Lee, Y. S. Meng, and S. Winkler, "A data-driven approach for accurate rainfall prediction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 9323–9331, 2019.

[26] H.-C. Kim, S. Pang, H.-M. Je, D. Kim, and S. Y. Bang, "Constructing support vector machine ensemble," *Pattern recognition*, vol. 36, no. 12, pp. 2757–2767, 2003.

[27] J. Du, Y. Liu, Y. Yu, and W. Yan, "A prediction of precipitation data based on support vector machine and particle swarm optimization (pso-svm) algorithms," *Algorithms*, vol. 10, no. 2, p. 57, 2017.

[28] K. Chau and C. Wu, "A hybrid model coupled with singular spectrum analysis for daily rainfall prediction," *Journal of Hydroinformatics*, vol. 12, no. 4, pp. 458–473, 2010.

[29] C.-W. Hsu, C.-C. Chang, C.-J. Lin *et al.*, "A practical guide to support vector classification," 2003.

[30] K. Duan, S. S. Keerthi, and A. N. Poo, "Evaluation of simple performance measures for tuning svm hyperparameters," *Neurocomputing*, vol. 51, pp. 41–59, 2003.

[31] S. Zainudin, D. S. Jasim, and A. A. Bakar, "Comparative analysis of data mining techniques for malaysian rainfall prediction," *Int. J. Adv. Sci. Eng. Inf. Technol*, vol. 6, no. 6, pp. 1148–1153, 2016.

[32] S. Cramer, M. Kampouridis, A. A. Freitas, and A. K. Alexandridis, "An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives," *Expert Systems with Applications*, vol. 85, pp. 169–181, 2017.

[33] M. Qiu, P. Zhao, K. Zhang, J. Huang, X. Shi, X. Wang, and W. Chu, "A short-term rainfall prediction model using multi-task convolutional neural networks," in *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2017, pp. 395–404.

[34] B. Klein, L. Wolf, and Y. Afek, "A dynamic convolutional layer for short range weather prediction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4840–4848.