



Article

Evaluation of Data Sufficiency for Interannual Knowledge Transfer of Crop Type Classification Models

Mohammadreza Osouli and Faramarz F. Samavati *

Department of Computer Science, University of Calgary, Calgary, AB T2N 1N4, Canada; mohammadreza.osouli@ucalgary.ca

* Correspondence: samavati@ucalgary.ca

Abstract: We present a study on the effectiveness of using varying data sizes to transfer crop type classification models from one year to the next, emphasizing the balance between data sufficiency and model accuracy. The significance of crop detection through satellite imaging lies in its potential to enhance agricultural productivity and resource management. Machine learning, particularly techniques like long short-term memory (LSTM) models, has become instrumental in interpreting these satellite data due to its predictive accuracy and adaptability. However, the direct application of models trained in one year to subsequent years poses challenges due to variations in environmental conditions and agricultural practices. Fine-tuning pre-existing models is a prevalent strategy to overcome these temporal discrepancies, though it necessitates a careful evaluation of the quantity and relevance of new data. This study explores the cost-benefit of fine-tuning existing models versus developing new ones based on the quantity of new data, utilizing LSTM models for their transferability and practicality in agricultural applications. Experiments conducted using satellite data from farms in southern Alberta reveal that smaller datasets, with fewer than 25 fields per class, can effectively fine-tune models for accurate interannual classification, while larger datasets are more conducive to training new models. This poses a key challenge in optimizing data usage for crop classification, straddling the line between data sufficiency and computational efficiency. The findings offer valuable insights for optimizing data use in crop classification, benefiting both academic research and practical agricultural applications.

**Citation:** Osouli, M.; F. Samavati, F.

Evaluation of Data Sufficiency for Interannual Knowledge Transfer of Crop Type Classification Models.

Remote Sens. **2024**, *16*, 2007.<https://doi.org/10.3390/rs16112007>

Academic Editors: Liang Sun and Lionel Bombrun

Received: 9 April 2024

Revised: 30 May 2024

Accepted: 31 May 2024

Published: 3 June 2024

Corrected: 2 July 2024

Keywords: crop type classification; data sufficiency; optimization; machine learning; neural networks; transfer learning

1. Introduction

The accurate mapping of crop types using satellite imagery is critical for enhancing global food security and managing agricultural resources efficiently. It provides essential insights for precision agriculture, enabling optimized crop management, yield prediction, and land use planning. The use of low-resolution satellite images from Sentinel-1, Sentinel-2, and Landsat satellites (with resolutions between 10 and 30 m) for crop classification is becoming a focal point in scholarly research. The images from these satellites have been widely used for training machine learning models [1]. A significant challenge in this field is the spatial dependency of data, which is affected by a range of factors like weather, soil quality, and elevation [2]. These variables are not consistent across different locations and can change from one year to the next, potentially reducing the effectiveness of a model trained on data from a previous year when applied to subsequent years [3]. Consequently, most research aims to either improve the accuracy of these crop classification models or to enhance their adaptability to diverse conditions, with an emphasis on maintaining spatial consistency and ensuring the models' relevance over multiple years [4,5].

One of the common steps in this model enhancement is labelling (i.e., identifying the types of crops grown on each farm field for the model). Since collecting these data through



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

field surveys or high-resolution images from drones or planes is costly, it is critical to make this process more efficient.

A common method to tackle the data labelling challenge is to fine-tune models trained on past data to match the current year's conditions using some new data from the current year. This technique has been examined in prior research [3,5,6], but a key detail is often overlooked: the amount of new data required to effectively and accurately update models for crop type classification. While larger new datasets can improve accuracy, it comes with an efficiency cost as gathering large datasets in this field of study requires large surveying or gathering high-resolution aerial imaging, which is cost-intensive labour. On the other hand, smaller-sized labelled datasets may result in low-accuracy models.

In this paper, as a stepping stone toward identifying optimal data size, we compare the different amounts of data needed to transfer a model to a new year while maintaining reliable accuracy. We also seek to find the point where it is more cost-effective to develop a new model instead of adjusting an existing one. Our goal is to specify the data quantity at which a new model surpasses the accuracy of an updated older model. As long short-term memory (LSTM) models are easy to transfer and are proven to be practical in this field [7,8], we primarily use this model.

Our approach serves as a valuable benchmark for determining the requisite amount of data for transferring crop type classification models, catering to both academic researchers and industry professionals.

For the completeness of our evaluation, we compare the result accuracy with a conventional random forest model trained using the same data. The random forest models have been used for a long time for crop type classification [2,9,10], so we used them as a comparison to assess the effectiveness and efficiency of our proposed method.

Our paper is structured as follows: In related works, we review existing research on crop type classification, setting the stage for our study. In Section 2, we detail our methodology, including data collection and the experimental setup. In Section 3, we present our experimental results, showcasing the performance of our models. In Section 4, we explain our results and discuss our findings. In Section 5, we summarize our findings and suggest areas for future research, particularly focusing on improving model transferability and data handling.

1.1. Related Works

In agricultural management and harvest prediction, identifying and mapping crop types is crucial. This process supplies important data to agencies involved in agriculture and insurance. If performed manually, documenting field details like crop type classification is expensive. Remote sensing technology offers a cost-effective solution, enabling the collection of a large amount of data at once and reducing labour costs [11]. However, on-site data are still needed to create and validate accurate classification models.

Recently, there has been a considerable amount of research work dedicated to crop detection using machine learning. These works can be categorized based on their emphasis on either specific features or specific models to use.

- Feature selection: These research works focus on determining the optimal combination of information to enhance model accuracy. They aim to identify which data inputs or features most significantly impact prediction outcomes.
- Model selection: These research works delve into experimenting with various models to identify those best suited for different scenarios. The emphasis in this area of study is on finding the most efficient and accurate algorithm or approach for crop classification.
- Studies on enhancing generality of the models: In the third category, there are works with a focus on solving specific problems in crop type classification such as the spatial transferability of the models [4].

In the subsequent three parts, we will review the related work according to the categories outlined above.

1.1.1. Feature Selection

Feature selection is a cornerstone in statistical machine learning. The right choice of features can enhance the performance of models [10]. This holds for crop type classification as well, where the selection of appropriate features is crucial for accurate identification.

A multitude of papers have been dedicated to this subject, seeking the most effective combination of satellite bands and other pertinent information [12,13]. For instance, one of the works by Sonobe et al. [10] investigated the crop classification potential of data from the Sentinel-1 and Sentinel-2 satellites for the year 2016. Their comprehensive approach included testing various sensors, bands, and models to evaluate the capability of these satellites in the realm of this task.

Another dimension of feature selection in crop classification is the amount of temporal information required for accurate results. The work by Vuolo et al. [14] delved into this aspect, comparing the effectiveness of single-image versus multi-temporal image datasets for classification tasks. They show that using multi-temporal images improves the accuracy of the model. In other work, Johnson and Mueller [15] introduced a method for large-scale early-stage crop classification and emphasized finding the earliest time that ensures an acceptable accuracy. Their method shows the importance of using previous data as ground truth to scale the result. Early-stage crop classification [16,17] has significant implications, especially for applications like food planning.

It has been shown by many that the combination of Sentinel-1 and Sentinel-2 imagery for crop classification can improve accuracy [18,19]. Notably, research by Van Tricht et al. [13] demonstrated that certain crops are more accurately classified using Sentinel-1 imagery than Sentinel-2.

Unlike other studies in this field that concentrate on the specific features used for model training, our work focuses on the quantity of data required to train or fine-tune a model. This approach makes our method versatile, allowing it to be applied to any model regardless of the features it utilizes.

1.1.2. Model Selection

As mentioned in the previous section, an example of such efforts is the study by Sonobe et al. [10]. Their work employed four distinct models—random forest (RF), support vector machine (SVM), multilayer feedforward neural networks, and kernel-based extreme learning machine (KELM)—to discern the most effective approach. They showed that the KELM model worked best for their study area. As research evolved, there was a noticeable inclination toward integrating neural network models into the crop classification paradigm. Zhao et al. [7] undertook an effort in this domain. Their study evaluated three deep learning architectures—1D convolutional neural network (1D CNN), long short-term memory (LSTM), and gated recurrent unit (GRU)—for early-stage crop classification, offering a comparative analysis of their efficacies. They showed that the LSTM model was the best choice for their study area among the three architectures.

Further broadening the horizon of model selection, Zhong et al. [8] presented a comprehensive comparison involving the XGBoost model and other models including multilayer perceptron, 1D CNN, LSTM, random forest, and SVM. They demonstrated that the best model based on overall accuracy was 1D CNN for their study area, but for specific crops, other models such as XGBOOST had better performance.

In a distinct exploration, Seydi et al. [20] delved into the potential of the dual attention CNN model for crop type classification, bringing a fresh perspective to the field. One of the more recent contributions by Wang et al. [21] tackled the challenge of unsupervised domain adaptation. Their innovative approach circumvented the issue of missing labels by adaptively harnessing information from alternate domains.

A major limitation of the previously mentioned studies on machine learning for crop detection is their reliance on data from a single year and a region with consistent conditions. Consequently, the models developed in these studies are only effective in the same year and

area where they were trained. The lack of generality in these models has been addressed in various ways by a third group of research works.

1.1.3. Studies on Enhancing Generality of the Models

While the domains of feature and model selection form the backbone of many research endeavours in crop type classification, numerous seminal works transcend these categorizations, addressing specific challenges inherent to the field.

One of the foremost challenges is the potential nontransferability of models across different geographical terrains. It is widely recognized that a model fine-tuned for one specific area might not deliver comparable accuracy when applied to another [4]. This gives rise to the substantial task of developing models robust enough to classify crops over expansive areas, such as entire countries or continents. The complexities involved in such endeavours are manifold, primarily due to the varied soil types, climatic conditions, and agricultural practices observed across large geographical expanses.

Song et al. [22] studied large-scale crop type classification. Their effort culminated in the introduction of a novel methodology tailored to classify soybeans across the vast expanse of the United States at a national scale. Such innovations have been mirrored in other regions as well, with notable research being carried out in Brazil [23] and Europe [24], further emphasizing the global importance and relevance of this challenge.

Temporal variation presents a unique obstacle in crop type classification. Specifically, the growth patterns of a crop can vary significantly from one year to another, influenced by a multitude of factors such as weather, soil quality, and other agronomic variables. This intrinsic variability implies that a model trained using data from one particular year might fail when applied to data from a subsequent year, even if the geographical region remains consistent [3].

This temporal transferability challenge remains relatively underexplored, with only a handful of research teams addressing it. Each has brought forth innovative solutions to mitigate the discrepancies caused by year-to-year variations.

A notable study in this domain was conducted by Bazzi et al. [6]. They ventured into the world of deep learning and employed the concept of knowledge distillation to address this challenge. The essence of their approach laid in creating two models, a Teacher model and a Student model. This framework allowed them to efficiently integrate new data from subsequent years with previously trained models, ensuring continuity and relevance.

Along this direction, Hao et al. [5] proposed a distinct approach, focusing on the random forest model. They harnessed the power of transfer learning and utilized the Crop-land Data Layer (CDL) as training samples. Their findings were promising, demonstrating that the transfer learning method exhibited superior accuracy compared to the standalone training of random forest models.

Extending the conversations from the preceding section, numerous studies have focused on crop type classification within the same year, typically training and applying models on data from a single annual cycle. Distinguishing our work, we target the applicability to industry needs by training our model on data from a preceding year and deploying it for predictions in the following year. We introduce a novel method and demonstrate its efficacy. Also, central to our approach is the goal of minimizing the amount of new year's data required. This goal has not been discussed by other related works such as [5,6]. Such data, typically obtained through surveys or high-resolution image classification techniques, is resource-intensive to gather. Our methodology seeks to optimize this process, ensuring accurate classification with a reduced data footprint.

2. Materials and Methods

Our objective is to determine the minimal amount of data necessary to adapt a machine learning model to a different year's model. For this purpose, we trained an LSTM model using data from 340 fields in southern Alberta from 2020, and tested its performance on both the same year's data and on data from 2021, which included 50 different fields

representing about 15% of our dataset. The initial results indicated that the model trained only on 2020 data performed poorly on the 2021 data, achieving only 58% accuracy across three classes.

To improve the model's adaptability, we fine-tuned it with varying amounts of 2021 data and evaluated its accuracy against a similar model that had no prior training. This process involved gathering batches of data with varying field counts per class for fine-tuning the 2020 model and measuring improvements in accuracy on the 2021 test data. For comparison, we also trained LSTM models from scratch with these same batches, and a random forest model to contrast with traditional methods. The LSTM model was selected for its proven suitability for such tasks, as validated by Zhao et al. [7]. A visual representation of this experimental setup is shown in Figure 1.

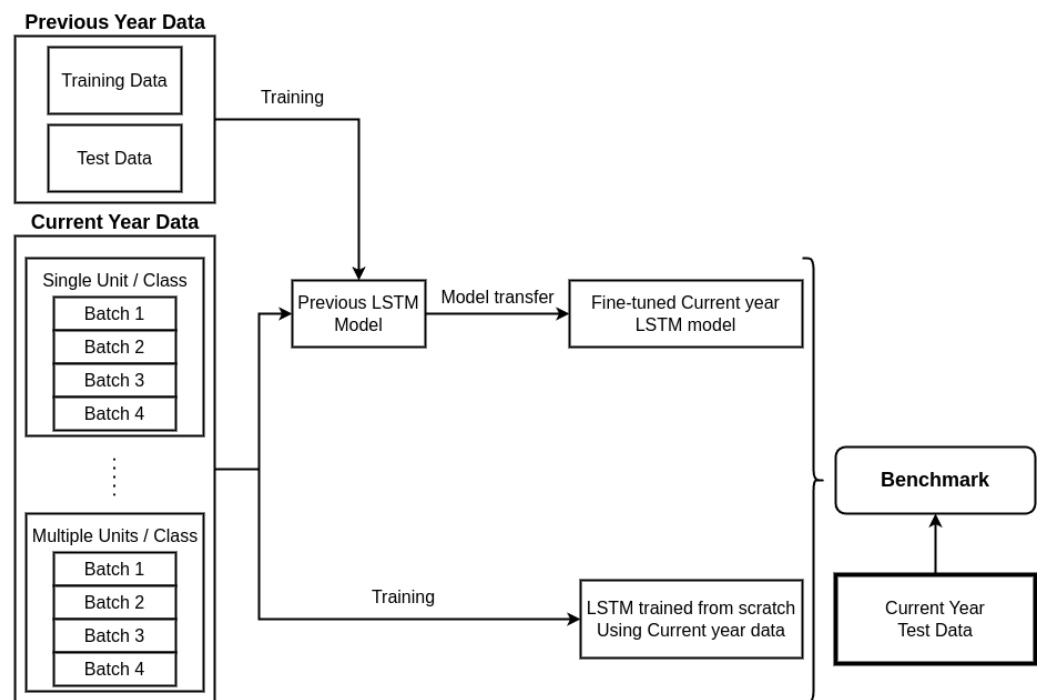


Figure 1. Flowchart of the process for evaluation of data sufficiency for interannual transfer of crop type classification models.

To conduct further experiments, we utilized the pretrained model of 2019 as the base for transferring to the years 2020 and 2021. We performed separate experiments for each year.

To be more precise, let us have the notation B_k for batches of the data for fine-tuning. The index k is the number of the fields in the batch where $k \in \{1, 5, 10, \dots\}$. We can denote the weights of the previous year's model as W_{pre} and W_{curr} for fine-tuned model weights.

$$W_{pre} \xrightarrow{\text{transfer}} W_{curr} \quad (1)$$

Our goal is to minimize the loss function L for different B_k values on a pretrained LSTM and compare its accuracy with the accuracy of a new LSTM that is trained using the same data. All accuracies are gathered using the test data, which has no intersection with training and validation data.

2.1. LSTM

Long short-term memory recurrent neural networks (LSTM RNNs) are sophisticated frameworks for processing sequences in neural network models [25]. These networks are ideal for crop classification as the pixel values of satellite images from different dates can be considered as time series. These models consist of two configurations: a many-to-

many setup, which produces outputs at each interval in the sequence, and a many-to-one arrangement, yielding an output solely at the sequence's end. Unlike conventional RNNs, LSTM networks can preserve information over extensive temporal intervals, addressing the traditional difficulties associated with gradient vanishing.

In the context of classification tasks, long short-term memory (LSTM) models excel by leveraging their ability to capture temporal dependencies in sequential data. To adapt an LSTM model for classification, the sequence of hidden states generated by the LSTM layers is typically followed by a fully connected (dense) layer that transforms the LSTM output to a dimensionality that matches the number of classes in the classification task. This transformation is crucial for mapping the rich, high-dimensional features learned by the LSTM to a more interpretable space relevant to the specific classification problem. The final layer of the model employs a softmax activation function, which converts the output of the fully connected layer into a probability distribution over the target classes. The softmax function ensures that the output values are non-negative and sum up to one, making them directly interpretable as class probabilities. This architectural arrangement allows the LSTM model to effectively process and classify sequential data, making informed predictions based on the temporal relationships it uncovers. Figure 2 shows a visual representation of this architecture.



Figure 2. Diagram of a simple LSTM classifier.

In the LSTM architecture, the state s_t at any given time t is determined by the following formalisms, also visualized in Figure 3:

- The input x_t is passed to the neuron;
- The hidden state h_t is derived by the formula $h_t = Ux_t + Ws_{t-1}$;
- The state s_t is updated by applying an activation function f to the hidden state, denoted as $s_t = f(Uh_t + b)$;
- The output y_t is then obtained via the transformation $y_t = g(Vs_t + c)$;

where U , V , and W are matrices that correspond to the weights of the network, while b and c serve as bias vectors. The functions f and g typically embody tanh and softmax activations, respectively.

Three distinct gates orchestrate the internal mechanics of the LSTM: the forget gate f_t , the input gate i_t , and the output gate o_t , which are illustrated in the following equations:

- Forget gate: $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$;
- Input gate: $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$;
- Cell candidate: $\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$;
- Update cell state: $C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$;
- Output gate: $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$;
- Update hidden state: $s_t = o_t \cdot \tanh(C_t)$.

Here, σ represents the sigmoid function, and the cell state C_t integrates the past cell state C_{t-1} with the current candidate state $\tilde{C}_t$, modulated by the forget gate f_t and input gate i_t , respectively. The output gate o_t subsequently governs the final state s_t by modulating the cell state with the tanh function.

The visuals in Figure 3 encapsulate the data flow within an LSTM cell, delineating how the gates selectively filter information throughout the network's iterative process.

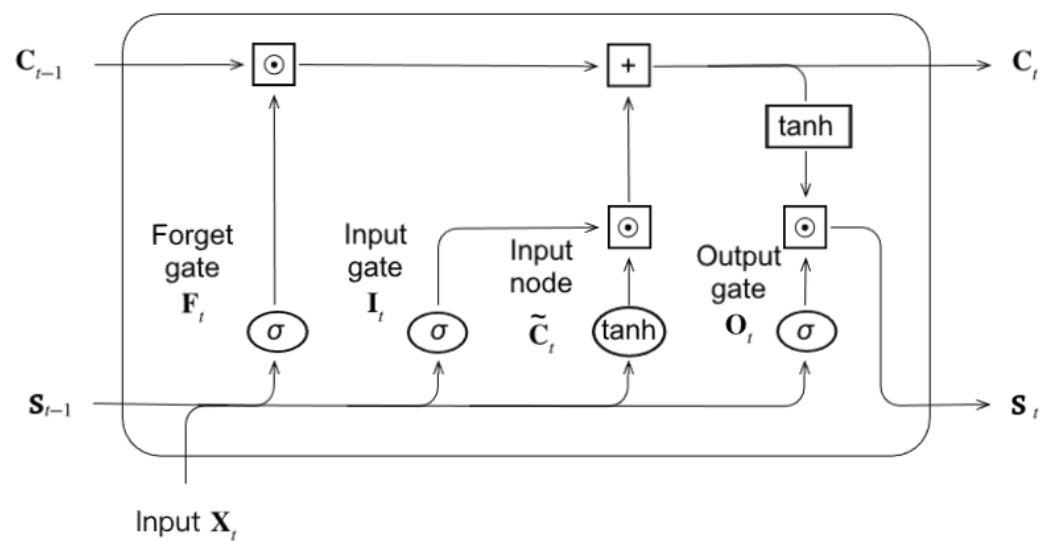


Figure 3. Diagram of a single LSTM neuron.

2.2. Random Forests

Random forests are an ensemble learning method used for classification and regression tasks that operate by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, it is the average prediction of the individual trees.

Each tree in the random forest is built from a sample drawn with replacement (bootstrap sample) from the training set. When splitting a node during the construction of the tree, the best split is chosen from a random subset of the features. This strategy of random feature selection when building trees ensures that the forest is diverse, which typically results in a more robust overall model.

A formal representation of the function employed by a random forest for a regression or classification problem might be given by the following:

$$Y = f(X) + \epsilon \quad (2)$$

where Y is the output, X represents the input vector, f signifies the model built by the random forest, and ϵ denotes the error term or the noise. The model f itself is the aggregated collection of decision trees f_i , each contributing a vote or a value:

$$f(X) = \frac{1}{B} \sum_{i=1}^B f_i(X) \quad (3)$$

for a regression problem, or by a majority vote for classification:

$$f(X) = \text{mode}\{f_1(X), f_2(X), \dots, f_B(X)\} \quad (4)$$

Here, B is the number of trees in the forest, and $f_i(X)$ is the prediction of the i -th tree. The individual tree predictions $f_i(X)$ are either class labels for classification problems or real numbers for regression problems. The final prediction Y is either the mean prediction (regression) or the mode of the predictions (classification) from all individual trees in the forest. A visual representation of the decision tree and random forest is provided in Figure 4.

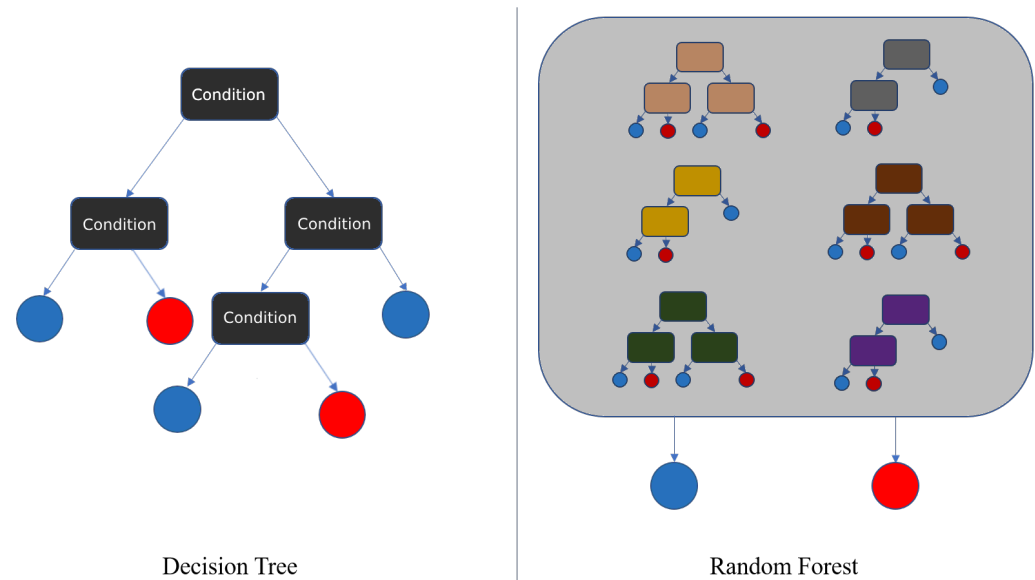


Figure 4. Decision tree and random forest structures. Different classes are visualized as blue and red circles.

2.3. Dataset

We opted for Sentinel-1 due to its widespread popularity and extensive use in numerous applications, including crop classification and environmental monitoring [10,12,13].

Sentinel-1 is a crucial satellite mission managed by the European Space Agency (ESA), consisting of a constellation of two polar-orbiting satellites. These satellites provide all-weather, day-and-night radar imaging, which is vital for a wide range of applications, from land and ocean monitoring to emergency response and disaster management.

Sentinel-1 operates in a conflict-free operation mode, completing a full cycle every 12 days. The satellite is outfitted with a C-band synthetic aperture radar (C-SAR). It offers four imaging modes: Strip Map (SM), Interferometric Wide-swath (IW), Extra Wide-swath (EW), and Wave (WV). Additionally, the C-SAR is capable of dual-polarization operation, supporting both HH (Horizontal Transmit/Horizontal Receive) + HV (Horizontal Transmit/Vertical Receive) and VV (Vertical Transmit/Vertical Receive) + VH (Vertical Transmit/Horizontal Receive) modes [26].

We choose to use the IW, which provides dual polarization of VV and VH. This means that every pixel in the dataset images has two bands with the same name of VV and VH, which represents the power of the radar signal that has been reflected back to the satellite after hitting the Earth's surface. These values are usually stored in digital numbers (DN) by quantizing the continuous analog raw signals into discrete digital values. Hereafter in this paper, the terms "VV" and "VH" will specifically refer to the pixel values measured in the VV and VH polarization bands of the Sentinel-1 imagery.

For our study, we gathered the Sentinel-1 images of 390 fields located in southern Alberta as shown in Figure 5. This dataset spans three years: 2019, 2020, and 2021. Out of these fields, 50 were randomly selected as the test dataset with a balanced distribution of classes. In these fields, each class has a share of around one-third of the total. Additionally, we randomly designated 20% of the remaining data as the validation set. This validation set is utilized to determine the optimal number of training epochs to avoid overfitting. We monitored our models' training procedure and stopped the training early when the performance on the validation set began to deteriorate while the training loss continued to decrease. Based on that, we used the epoch count of 10. After that, the entire dataset, comprising the rest of the 340 fields, was employed for further training and fine-tuning of the models.

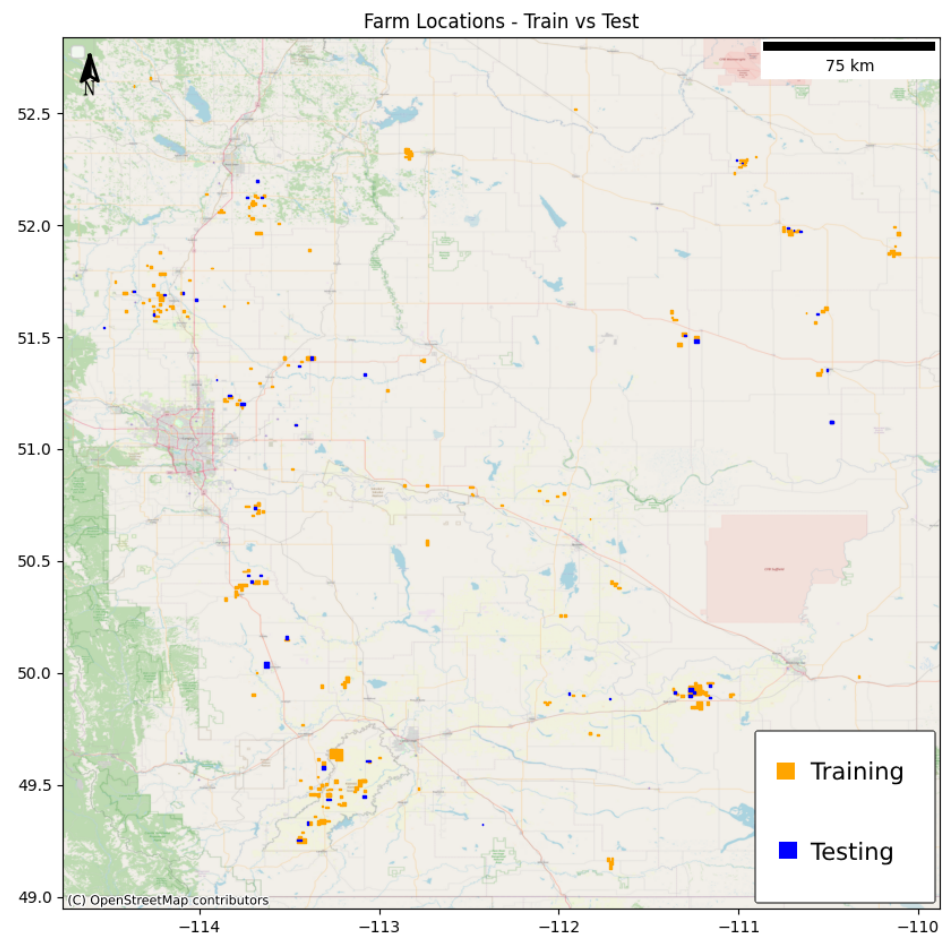


Figure 5. Data fields in Alberta. The orange boxes are training and validation data, and the blue boxes are testing data.

The number of Sentinel-1 images available for the years 2019, 2020, and 2021 are varying, spanning from 1 March to 1 September. This timeframe covers the entire farming season in Alberta, dictated by the region's climatic conditions. Available dates of Sentinel-1 images are listed in Table 1.

As the ground truth, we leverage the Agriculture and Agri-Food Canada (AAFC) [27] dataset, which provides a comprehensive crop inventory. This dataset, collected by the Government of Canada, is made publicly available on an annual basis and covers the whole agricultural area of Canada. It contains more than 10 different crop types being farmed in Canada. The data are presented as GeoTIFF images with a resolution of 30 m per pixel, as demonstrated in Figure 6. In this work, we chose the three most-planted crops in south Alberta—canola, barley, and spring wheat—as they cover over 90% of the agricultural areas in this province.

Table 1. Data availability for the area for years 2019, 2020, and 2021.

Year	Days of the Year
2019	79, 82, 91, 94, 103, 106, 115, 118, 127, 130, 135, 139, 142, 147, 151, 154, 159, 163, 166, 171, 175, 178, 183, 187, 190, 195, 199, 202, 207, 211, 214, 219, 223, 226, 231, 235, 238
2020	73, 77, 85, 89, 97, 101, 109, 113, 121, 125, 133, 137, 145, 149, 157, 161, 169, 173, 181, 193, 197, 205, 209, 217, 221, 229, 233, 241
2021	79, 80, 83, 92, 95, 104, 107, 116, 119, 128, 131, 140, 143, 152, 155, 164, 167, 176, 179, 188, 191, 200, 203, 212, 215, 224, 227, 236, 239

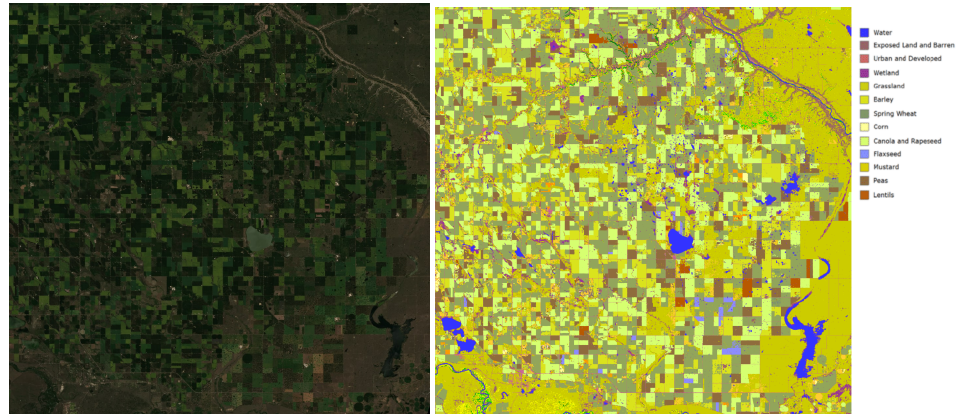


Figure 6. Image of an area located in south Alberta and its corresponding Agriculture and Agri-Food Canada (AAFC) crop tags. Each colour represents a crop or a nonagricultural area such as wetlands or urban areas.

2.4. Sampling

To create a representative sample for both training and testing our models, we employ a sampling strategy that focuses on the central areas of each farm field, which is generally more consistent in terms of crop type presence and avoids the noisier boundary regions where edge effects can occur.

Because our fields have almost the same amount of area, we need to have the same number of data points from each to have an unbiased dataset. On one hand, having a low number of data points will cause not enough variations of data points in different conditions, such as different elevations or soil conditions. On the other hand, having a lot of data points will cause redundancy of data, which can lead to overfitting. Thus, we standardized our approach by sampling 20 pixels from each field. These pixels are chosen randomly, but with a probability density that decreases as the distance from the centre of the field increases. This strategy helps to minimize the inclusion of outlier data that may not be representative of the main crop type within the field. Figure 7 represents a visualization of the sampling method.



Figure 7. Visualization of sampling method for a field. The black star is the centre of the field and the red stars are the random samples in the area of distance d .

2.5. Features

Since the Sentinel-1 satellite imagery is not captured on consistent dates across different years, using a straightforward temporal array as our feature is not feasible. This is because the same index in a feature array for one year may not correspond to the same calendar day in the array for another year. To address this issue, we developed a uniform feature format that can accommodate data that spans multiple years.

Given the variable time intervals between the Sentinel-1 satellite images, it becomes necessary to fill in missing data to create a consistent model applicable to both years. To achieve this, we employed a linear interpolation method. While other interpolation methods could potentially impact the model's accuracy, examining these alternatives is out of the scope of this work.

For every pixel and band (either VV or VH), we utilized an array comprising 365 elements, each representing a day of the year, and populated it with pixel values corresponding to specific days. Given that the revisit time of Sentinel-1 is more than six days, it is not possible to have images of the same field in two consecutive days. Therefore, to obtain more compact features, we halved the array's size. Consequently, every array index i accounts for two days: $2i$ and $2i + 1$.

Also, as the models are sensitive to missing values and we want to use one model for multiple years, we need to interpolate the values for missing data. The problem comes from the fact that the visiting days of the Sentinel-1 satellite are different based on days as it is available in Table 1. To standardize the data for different years, we linearly interpolated the values between the days to create a consistent representation. One example of available data vs. interpolated feature is visualized in Figure 8. Additionally, a visual representation of the process is provided in Figure 9.

Recognizing that farming activities in Alberta are dormant from January through March and September to December, we eliminated these periods from our feature set to further streamline its size. Therefore, the index 0 of the feature array corresponds to the 72nd day of the year (13 March) and the last index is for the 246th day of the year (3 September).

The feature array has a size of 88 (see Figure 8), and as there are two different bands, the dimensions of the final tensor being passed to the model are 88 by 2.

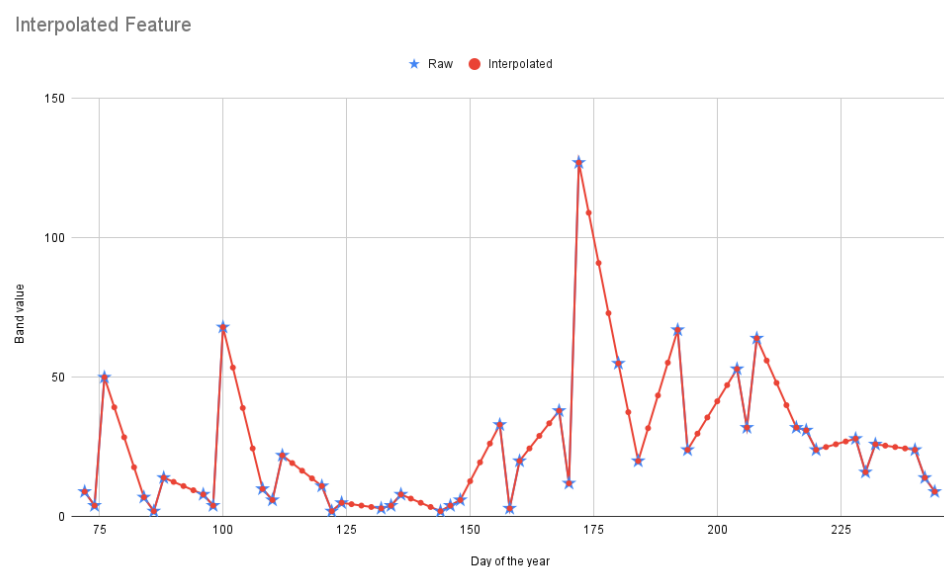


Figure 8. The data points with star marks are the available data, while the red circles are interpolated values to have a consistent feature structure for the models.

Index	1	2		...	...		87	88
DOY	72	74	76	...	...	242	244	246
Band	0.0	null	0.4	..	..	0.2	null	0.0

↓

DOY	72	74	76	...	...	242	244	246
Band	0.0	0.02	0.4	..	..	0.2	0.1	0.0

Figure 9. **Top:** The raw data of VV (Vertical/Vertical). **Bottom:** The array after interpolation. The value of zero is used for the start and end missing data to have a proper interpolation.

Also, we created varying batch sizes to assess the impact on the performance of fine-tuned pretrained model. Specifically, we employed batches containing 1, 5, 10, and incrementally up to 40 fields per class, enabling a comprehensive evaluation of model adaptability and efficiency under different data scales.

2.6. Models

Our primary objective was to determine the optimal amount of data required to facilitate the transferability of models across different years. It is crucial, therefore, to employ a model inherently adept at such transfers. Given that artificial neural networks (ANNs) are well regarded for their transferability capabilities [28], and considering the proven efficacy of long short-term memory (LSTM) models in crop classification tasks [7], we chose to leverage LSTM for our experiments.

For a comprehensive evaluation, we compared the outcomes of the transferred model against results from the freshly trained LSTM model. This comparison was conducted using varying amounts of transferred data to provide insights into the efficacy and efficiency of each approach.

For our LSTM-based experiments, we adopted the same architecture as proposed by Zhao et al. [7]. They used the same satellite with the same format of missing data and the number of classes they have is five, which is almost the same as our number of classes. The other reason that we use the same model is that they achieved high accuracy in early-stage crop type classification.

Our model comprises three LSTM layers, each housing 100 units, culminating in a softmax output layer. The number of layers and units is the result of trying different structures and choosing the best one due to the structure of the data [7]. A visual representation of the model structure is available in Figure 10.

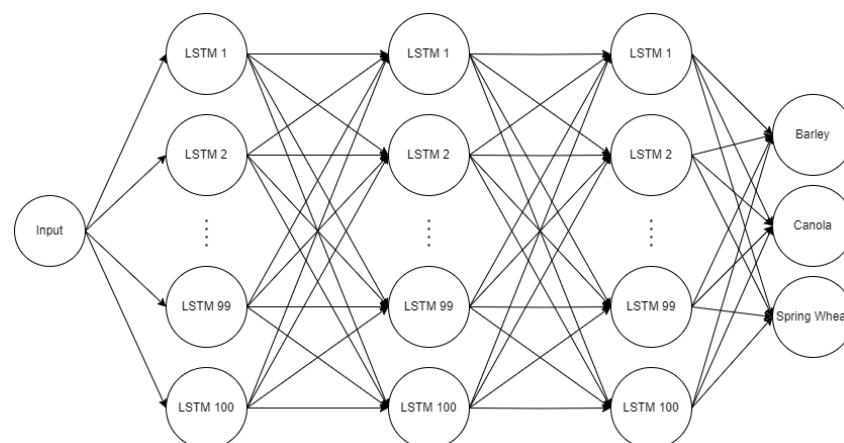


Figure 10. Three-layer long short-term memory (LSTM) model.

2.7. Training Procedure

Our primary objective in the first experiment was to adopt a 2020-based model for 2021 crop classification, focusing on predominant crops in Alberta: barley, canola, and spring wheat.

The initial step involved training an LSTM model using the complete 2020 dataset, serving as our foundational model. For the 2021 model adaptation, we prepared distinct batches of data. The initial batch contained data from just one field for each of the three crop types, totalling three fields. Subsequent batches progressively expanded, with each encompassing an additional five fields per crop type, culminating in batches of 1, 5, 10, 15, 20, 25, 30, 35, and 40 fields for every crop. To overcome the bias in data, we repeated the experience four times with different randomization of the fields.

To fine-tune the 2020 model, we kept its original structure and parameters to preserve the learned features and retrained the model on a new batch of data. To achieve this, we employed a reduced learning rate of 0.0001, which is considerably lower than the initial rate of 0.001.

Using these 2021 data batches, we then fine-tuned the 2020 LSTM model. For a comprehensive evaluation, we also trained both an LSTM and a random forest model from scratch using these same batches. Ultimately, we gauged the efficacy of these models using the test dataset. The ensuing results and our conclusions are delineated in Section 3.

To ensure the reliability of our findings, we repeated the experiment two more times. We fine-tuned a 2019 model using data from both 2020 and 2021 for the same region. We then compared these outcomes with models that were independently trained using 2020 and 2021 data. For each year, we also included two random forest models for comparative analysis.

2.8. Implementation

To implement the experience, we used the Python programming language version 3.10. The data were gathered from the Sentinel Hub website. We used the Keras library version 2.15.0 for LSTM training with the backend of Tensorflow version 2.15.0. Also, the scikit-learn version 1.2.2 was used for training the RF model. For easier maintenance, all the parts were implemented in the Google Colab [29] environment. Training models from scratch on a parallel system equipped with graphics processing units (GPU) requires approximately 30 min. In contrast, fine-tuning a pretrained model on the same system takes roughly 10 min.

3. Results

In this section, we provide a detailed presentation of the experimental results obtained from our study. The performance outcomes, reflecting the efficacy of various modelling approaches, including fine-tuning and training from scratch, are presented in this section.

3.1. Overall Performance of the Models

Table 2 details the performance outcomes for the LSTM model by fine-tuning 2020 model for 2021 using different numbers of fields, ranging from 1 to 40 fields per class. The table's initial row specifically highlights the 2020 and 2021 test data performance of the pretrained model prior to any fine-tuning. We use the weighted F1-score average metric to show the performance of the models. For each number of fields per class, we fine-tuned the pretrained models four times using different randomly selected fields. We use the average of all four experiments' F1-score as a final metric for comparison.

A key point of interest is the baseline performance of the original model developed using 2020 data; when this model is applied to the 2021 test dataset without any fine-tuning, it achieves a weighted F1-score average of 0.58. This initial outcome sets a foundational benchmark for evaluating the subsequent improvements achieved through our fine-tuning processes and highlights the intrinsic variability and adaptability challenges associated with annual crop type classification. For example, Table 2 shows that if we use 15 fields per

class, for 4 experiments of selecting 15 fields each time at random, on average we expect to have an F1-score of 0.65. A graphical representation comparing these results is illustrated in Figure 11.

In the other experiments, we fine-tuned the 2019 model for the years 2020 and 2021. The initial F1-score of the 2019 model on these years are, respectively, 0.61 and 0.58. The accuracy for various fields per class counts is shown in Tables 3 and 4. A graphical summary is shown in Figure 12.

Table 2. Fine-tuned LSTMs weighted F1-score (2020 $\xrightarrow{\text{Transfer}}$ 2021).

2021 Fine-Tuned LSTM: Base Model 2020						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
No fine-tuning	0%	58	-	-	-	58
1	2.50%	43	56	52	62	53.25
5	12.50%	53	65	57	64	59.75
10	25%	59	64	62	64	62.25
15	37.50%	68	65	65	64	65.50
20	50%	65	65	62	64	64.00
25	62.50%	70	65	64	66	66.25
30	75%	66	65	67	63	65.25
35	87.50%	62	62	67	68	64.75
40	100%	64	62	65	73	66.00

Table 3. Fine-tuned LSTMs weighted F1-score (2019 $\xrightarrow{\text{Transfer}}$ 2021).

2021 Fine-Tuned LSTM: Base Model 2019						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
No fine-tuning	0%	58	-	-	-	58
1	2.50%	61	54	64	38	54.25
5	12.50%	63	61	62	46	58.00
10	25%	59	60	67	67	63.25
15	37.50%	60	59	70	69	64.50
20	50%	63	65	72	70	67.50
25	62.50%	63	71	72	69	68.75
30	75%	73	70	69	67	69.75
35	87.50%	71	72	69	68	70.00
40	100%	67	71	68	72	69.50

Table 4. Fine-tuned LSTMs weighted F1-score (2019 $\xrightarrow{\text{Transfer}}$ 2020).

2020 Fine-Tuned LSTM: Base Model 2019						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
No fine-tuning	0%	61	-	-	-	61
1	2.50%	52	78	71	62	65.75
5	12.50%	63	69	68	59	64.75
10	25%	71	71	59	66	66.75
15	37.50%	69	68	60	63	65.00
20	50%	69	66	68	67	67.50
25	62.50%	66	68	65	67	66.50
30	75%	67	70	65	75	69.25
35	87.50%	70	70	67	76	70.75
40	100%	73	73	72	77	73.75

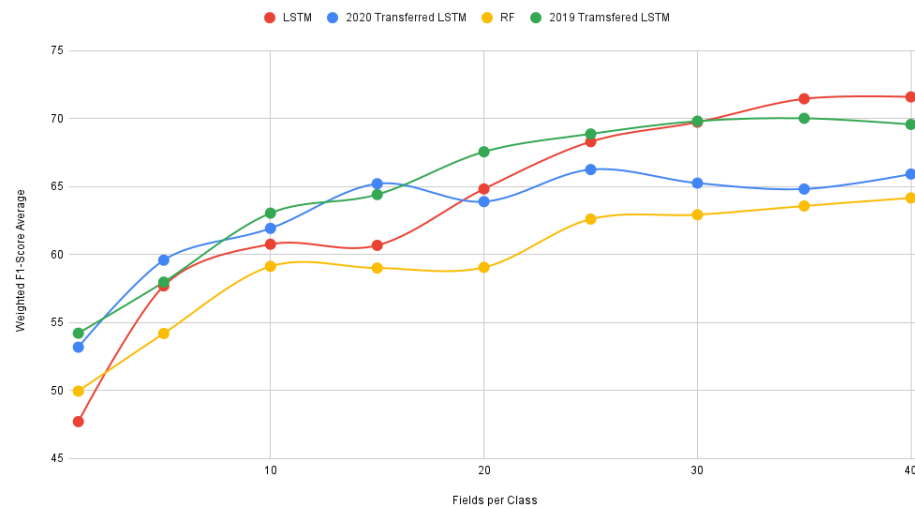


Figure 11. Weighted F1-score average of 2021 LSTM models transferred from 2020 and 2019 versus LSTM model trained from scratch and RF trained from scratch using 2021 data.

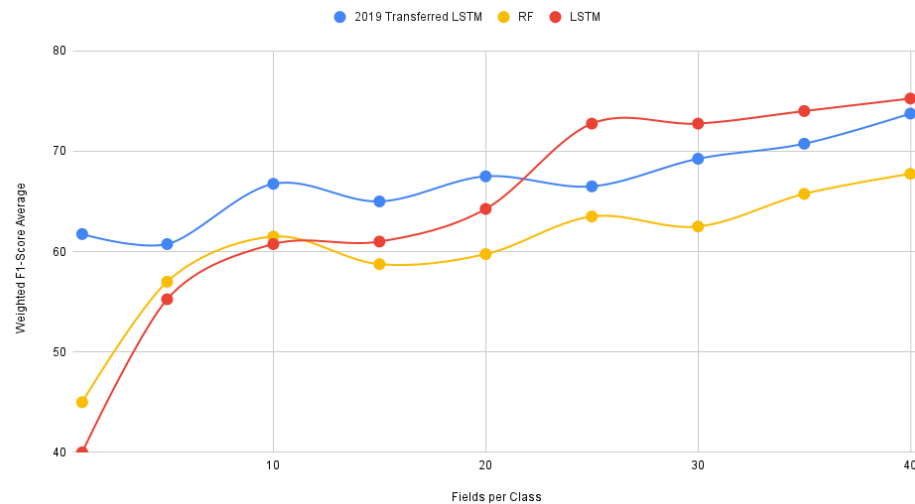


Figure 12. Weighted F1-score average of 2020 LSTM model transferred from 2019 versus LSTM model trained from scratch and RF trained from scratch using 2020 data.

Table 5 shows the weighted F1-score average results of LSTM models that were trained from scratch using 2020 and 2021 data. This training dataset is identical to the one utilized for the fine-tuning process.

Table 5. Overall accuracy table of LSTM method.

LSTM Trained from Scratch Using 2021 Data						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
1	2.50%	46	42	46	56	47.50
5	12.50%	53	61	57	59	57.50
10	25%	58	64	61	60	60.75
15	37.50%	59	63	63	57	60.50
20	50%	64	64	64	67	64.75
25	62.50%	68	69	70	67	68.50
30	75%	65	72	69	73	69.75
35	87.50%	70	70	73	73	71.50
40	100%	69	73	73	72	71.75

Table 5. Cont.

LSTM Trained from Scratch Using the 2020 Data						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
1	2.50%	52	78	71	62	65.75
5	12.50%	63	69	68	59	64.75
10	25%	71	71	59	66	66.75
15	37.50%	69	68	60	63	65.00
20	50%	69	66	68	67	67.50
25	62.50%	66	68	65	67	66.50
30	75%	67	70	65	75	69.25
35	87.50%	70	70	67	76	70.75
40	100%	73	73	72	77	73.75

3.2. Performance of the Models on Each Crop

Additionally, we evaluated the performance of each model on individual crop classes. Our findings indicate that barley and spring wheat exhibited similar behaviour, achieving higher F1-scores with fewer data in the fine-tuned LSTM model, and they achieved better scores in the LSTM trained from scratch with larger datasets, as illustrated in Figure 13. Conversely, canola displayed a reverse trend, indicating that increased data training leads to overfitting for this particular crop in the model, thereby reducing its overall accuracy. Such patterns suggest potential improvements through the adoption of crop-specific binary models.

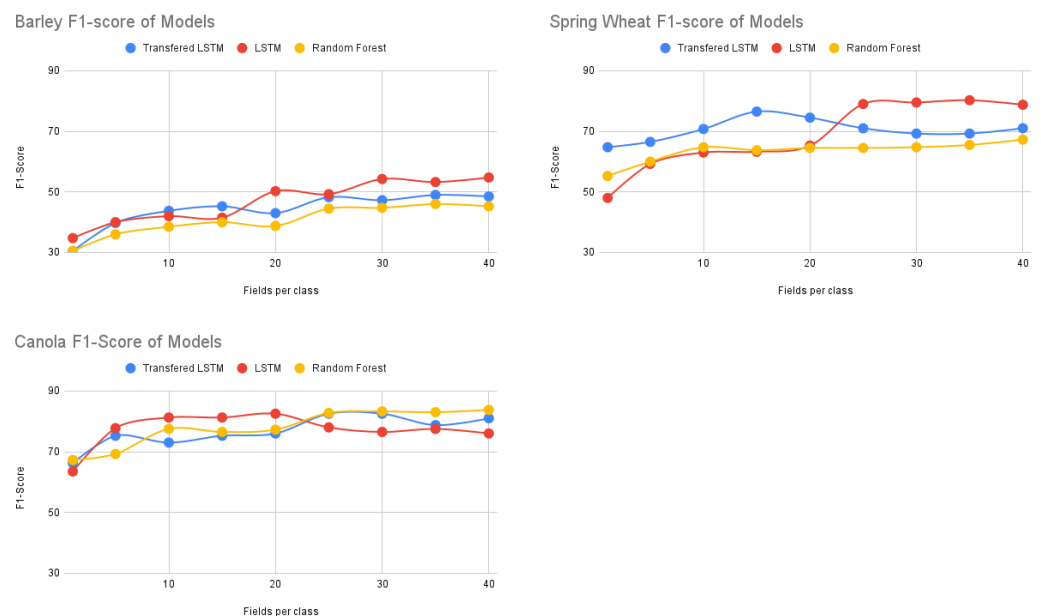


Figure 13. Comparison of F1-Score of different crops in the experiment (2020 $\xrightarrow{\text{Transfer}}$ 2021).

4. Discussion

Our research results show the effectiveness of fine-tuning LSTM models compared to training new models from scratch for annual crop type classification. The results indicate that fine-tuning leads to higher accuracy when fewer data points exist. Expectedly, we found that a fully-trained new model outperforms a fine-tuned model when a larger dataset is available. For example, comparing the transferred model of 2021 from 2020 versus training from scratch for 2021 (see Figure 11) indicates that when the data are smaller than 20 fields per class (50%) the transferred model maintains higher accuracy levels. In contrast, larger datasets support the development of new models with better

performance metrics. Our results for other experiments are generally consistent with this observation (see Tables 2–6). Based on these results, using smaller new datasets for fine-tuning is preferred to reduce the cost of surveying and gathering new data. Our results offer a comprehensive understanding of the trade-off between using smaller datasets for per-class fine-tuning and the resulting accuracy. In comparison, Bazzi et al. [6] used batch sizes of 60, 120, 180, and 240 for their training, which corresponded to an average of 30, 60, and 120 samples per class as they had two classes of irrigated and nonirrigated areas. They stated that for 30 samples per class, the model trained from scratch could not compete with the transferred models. In contrast, we demonstrated that for LSTM models, 30 samples per class are sufficient for training a model from scratch. This difference may arise from the structural differences between the models and the greater flexibility of LSTMs for this task. Unfortunately, they did not provide an analysis for smaller batch sizes.

Table 6. Overall accuracy table of random forest models.

RF Trained from Scratch Using the 2021 Data						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
1	2.50%	52	48	47	54	50.25
5	12.50%	52	58	53	54	54.25
10	25%	55	62	66	53	59.00
15	37.50%	59	61	61	56	59.25
20	50%	54	63	62	57	59.00
25	62.50%	65	64	61	61	62.75
30	75%	65	62	63	61	62.75
35	87.50%	64	62	63	66	63.75
40	100%	63	63	63	67	64.00
RF Trained from Scratch Using the 2020 Data						
Field per Class	Data Percentage	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average
1	2.50%	30	65	43	42	45.00
5	12.50%	45	67	57	59	57.00
10	25%	65	70	55	56	61.50
15	37.50%	63	63	51	58	58.75
20	50%	57	61	57	64	59.75
25	62.50%	58	61	66	69	63.50
30	75%	62	58	63	67	62.50
35	87.50%	65	62	64	72	65.75
40	100%	68	67	66	70	67.75

Furthermore, the results of our study indicate that the accuracy of the 2021 model transferred from 2019 is higher than that of the model transferred from 2020 (as depicted in Figure 11). This implies that the base model for fine-tuning does not necessarily need to be from the immediately preceding year. Hu et al. [3] confirmed the same conclusion by training a model from 2020 to 2019. In their method, they fine-tuned RF models using 3741 samples from 2019 for 5 classes. The main model of 2020 was pretrained using 5225 samples. This also highlights the larger amount of data that RF models require for fine-tuning. In comparison, we demonstrate that LSTM models need significantly less data for fine-tuning. In another study by Hao et al. [5], they transferred RF models using 500 training samples for a region in Alberta for two classes, spring wheat and canola, achieving an overall accuracy of 86%. This also demonstrates the substantial amount of data required for RF models to achieve acceptable accuracy.

Identifying the most appropriate year (or years) for knowledge transfer is a promising research area, especially when there is a more extensive history of remote sensing data. As an insight into this challenge, fine-tuning with smaller datasets is expected to be more effective and adaptable when there are minor changes in environmental and agricultural practices from year to year. This assumption is based on the idea that LSTM models can

adapt to new data without the need for extensive retraining as long as the new data do not significantly differ from the original training data in their fundamental characteristics.

Results also indicated that although the random forest model, such as [3,5,7], generally underperforms compared to the LSTM model, it may perform slightly better with small data volumes.

These findings highlight the need to evaluate data adequacy and model selection based on the specific accuracy requirements and available resources. Practically, this means that stakeholders in agricultural data analytics must carefully weigh the cost of data collection against the benefits of improved model performance through more extensive training or fine-tuning.

5. Conclusions

In this paper, we investigated the effectiveness of fine-tuning versus training LSTM models from scratch for crop classification, using several experiments. We fine-tuned several LSTM models using varying amounts of data, starting with models pretrained on 2019 and 2020 data, and then incorporating 2021 data. Additionally, to evaluate a different fine-tuning scenario, we conducted experiments by fine-tuning a model pretrained on 2019 data with 2020 data. In all experiments, we compared the results against LSTM and random forest (RF) models that were trained from scratch. Notably, the advantage of the transferred LSTM models over random forest models became less distinct as data volume increased after the threshold of 15 fields per class. Our analysis also revealed crop-specific trends, suggesting the potential for tailored approaches. This study emphasizes the importance of dataset size and specificity in optimizing model performance for agricultural applications.

While our research did not engage with feature transfer techniques, they represent a noteworthy direction for future exploration. Feature transfer often modifies only the final layers of a network, potentially preserving the majority of the pretrained model's structure. This method may not induce sufficient model adaptation, particularly in regions like southern Alberta, where crops exhibit closely aligned growth patterns and may thus require more extensive retraining of the model to capture the subtle distinctions necessary for accurate classification. Future studies could investigate the viability of feature transfer in such nuanced agricultural settings. Also, there are several more avenues to explore. One possibility is to investigate other machine learning models and compare their fine-tuning capabilities with LSTMs. Another direction for extension could involve early-stage crop detection, incorporating a time-threshold parameter to introduce an additional dimension to the study. Also, the investigation of finding a method to determine the best base model for knowledge transfer may be a good area to explore.

Author Contributions: Conceptualization, M.O. and F.F.S.; methodology, M.O. and F.F.S.; software, M.O.; validation, M.O. and F.F.S.; formal analysis, M.O. and F.F.S.; investigation, M.O.; resources, M.O. and F.F.S.; data curation, M.O.; writing—original draft preparation, M.O.; writing—review and editing, M.O. and F.F.S.; visualization, M.O.; supervision, F.F.S.; project administration, F.F.S.; funding acquisition, F.F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Mitacs under grant number “MITACS, IT32167”.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: We wish to acknowledge the support of the Mathematics of Information Technology and Complex Systems (MITACS), Decisive Farming (TELUS Agriculture) and the Natural Sciences and Engineering Research Council of Canada (NSERC). We would like to express our sincere thanks to Vincent Yeow Chieh Pang and Garret Duffy for the great discussions and insights during the Mitacs project. Special thanks go out to Christopher Mossman for proofreading the manuscript. We thank all members of the GIV research team at the University of Calgary, specifically Meysam Kazemi for helping in the implementation of the infrastructure of satellite imagery downloading, and all other members for their insightful discussions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Aiyim Orynbaikyzy, U.G.; Conrad, C. Crop type classification using a combination of optical and radar remote sensing data: A review. *Int. J. Remote Sens.* **2019**, *40*, 6553–6595. [\[CrossRef\]](#)
2. Tran, K.H.; Zhang, H.K.; McMaine, J.T.; Zhang, X.; Luo, D. 10 m crop type mapping using Sentinel-2 reflectance and 30 m cropland data layer product. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102692. [\[CrossRef\]](#)
3. Hu, Y.; Zeng, H.; Tian, F.; Zhang, M.; Wu, B.; Gilliams, S.; Li, S.; Li, Y.; Lu, Y.; Yang, H. An Interannual Transfer Learning Approach for Crop Classification in the Hetao Irrigation District, China. *Remote Sens.* **2022**, *14*, 1208. [\[CrossRef\]](#)
4. Orynbaikyzy, A.; Gessner, U.; Conrad, C. Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2. *Remote Sens.* **2022**, *14*, 1493. [\[CrossRef\]](#)
5. Hao, P.; Di, L.; Zhang, C.; Guo, L. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples. *Sci. Total Environ.* **2020**, *733*, 138869. [\[CrossRef\]](#)
6. Bazzi, H.; Ienco, D.; Baghdadi, N.; Zribi, M.; Demarez, V. Distilling Before Refine: Spatio-Temporal Transfer Learning for Mapping Irrigated Areas Using Sentinel-1 Time Series. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1909–1913. [\[CrossRef\]](#)
7. Zhao, H.; Chen, Z.; Jiang, H.; Jing, W.; Sun, L.; Feng, M. Evaluation of Three Deep Learning Models for Early Crop Classification Using Sentinel-1A Imagery Time Series—A Case Study in Zhanjiang, China. *Remote Sens.* **2019**, *11*, 2673. [\[CrossRef\]](#)
8. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [\[CrossRef\]](#)
9. Feng, S.; Zhao, J.; Liu, T.; Zhang, H.; Zhang, Z.; Guo, X. Crop Type Identification and Mapping Using Machine Learning Algorithms and Sentinel-2 Time Series Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3295–3306. [\[CrossRef\]](#)
10. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K.i. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. *Gisci. Remote Sens.* **2017**, *54*, 918–938. [\[CrossRef\]](#)
11. Ryu, C.; Suguri, M.; Iida, M.; Umeda, M.; Lee, C. Integrating Remote Sensing and GIS for prediction of rice protein contents. *Precis. Agric.* **2010**, *12*, 378–394. [\[CrossRef\]](#)
12. Adrian, J.; Sagan, V.; Maimaitijiang, M. Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 215–235. [\[CrossRef\]](#)
13. Van Tricht, K.; Gobin, A.; Gilliams, S.; Piccard, I. Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping: A Case Study for Belgium. *Remote Sens.* **2018**, *10*, 1642. [\[CrossRef\]](#)
14. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.T. How much does multi-temporal Sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 122–130. [\[CrossRef\]](#)
15. Johnson, D.M.; Mueller, R. Pre-and within-season crop type classification trained with archival land cover information. *Remote Sens. Environ.* **2021**, *264*, 112576. [\[CrossRef\]](#)
16. Xing, H.; Chen, B.; Lu, M. A Sub-Seasonal Crop Information Identification Framework for Crop Rotation Mapping in Smallholder Farming Areas with Time Series Sentinel-2 Imagery. *Remote Sens.* **2022**, *14*, 6280. [\[CrossRef\]](#)
17. Zhang, H.; Kang, J.; Xu, X.; Zhang, L. Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China. *Comput. Electron. Agric.* **2020**, *176*, 105618. [\[CrossRef\]](#)
18. Prins, A.J.; Van Niekerk, A. Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. *Geo-Spat. Inf. Sci.* **2021**, *24*, 215–227. [\[CrossRef\]](#)
19. Blickensdörfer, L.; Schwieder, M.; Pflugmacher, D.; Nendel, C.; Erasmi, S.; Hostert, P. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sens. Environ.* **2022**, *269*, 112831. [\[CrossRef\]](#)
20. Seydi, S.T.; Amani, M.; Ghorbanian, A. A Dual Attention Convolutional Neural Network for Crop Classification Using Time-Series Sentinel-2 Imagery. *Remote Sens.* **2022**, *14*, 498. [\[CrossRef\]](#)
21. Wang, Y.; Feng, L.; Zhang, Z.; Tian, F. An unsupervised domain adaptation deep learning method for spatial and temporal transferable crop type mapping using Sentinel-2 imagery. *ISPRS J. Photogramm. Remote Sens.* **2023**, *199*, 102–117. [\[CrossRef\]](#)
22. Song, X.P.; Potapov, P.V.; Krylov, A.; King, L.; Di Bella, C.M.; Hudson, A.; Khan, A.; Adusei, B.; Stehman, S.V.; Hansen, M.C. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. *Remote Sens. Environ.* **2017**, *190*, 383–395. [\[CrossRef\]](#)
23. Pott, L.P.; Amado, T.J.C.; Schwalbert, R.A.; Corassa, G.M.; Ciampitti, I.A. Satellite-based data fusion crop type classification and mapping in Rio Grande do Sul, Brazil. *ISPRS J. Photogramm. Remote Sens.* **2021**, *176*, 196–210. [\[CrossRef\]](#)
24. d'Andrimont, R.; Verhegghen, A.; Lemoine, G.; Kempeneers, P.; Meroni, M.; van der Velde, M. From parcel to continental scale—A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. *Remote Sens. Environ.* **2021**, *266*, 112708. [\[CrossRef\]](#)
25. Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to Forget: Continual Prediction with LSTM. *Neural Comput.* **2000**, *12*, 2451–2471. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Torres, R.; Snoeij, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.; Rommen, B.; Floury, N.; Brown, M.; et al. GMES Sentinel-1 mission. *Remote Sens. Environ.* **2012**, *120*, 9–24. [\[CrossRef\]](#)
27. Fisette, T.; Rollin, P.; Aly, Z.; Campbell, L.; Daneshfar, B.; Filyer, P.; Smith, A.; Davidson, A.; Shang, J.; Jarvis, I. AAFC annual crop inventory. In Proceedings of the 2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Fairfax, VA, USA, 12–16 August 2013; pp. 270–274. [\[CrossRef\]](#)

-
28. Bozinovski, S. Reminder of the First Paper on Transfer Learning in Neural Networks, 1976. *Informatica* **2020**, *44*, 291–302. [[CrossRef](#)]
 29. Google Colaboratory. Available online: <https://colab.research.google.com/> (accessed on 8 June 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.