Multimodal Fusion of Sentinel 1 images and Social Media Data for Snow Depth estimation

Damianos Mantsis, Marios Bakratsas, Stelios Andreadis, Petteri Karsisto, Anastasia Moumtzidou, Ilias Gialampoukidis, Ari Karppinen, Stefanos Vrochidis, Ioannis Kompatsiaris, *Senior Member, IEEE*

Abstract—Recent developments in remote sensing have shown that snow depth can be estimated accurately on a global scale using satellite images through cross-polarization and co-polarization backscatter measurements. This method does however have some limitations in low-land areas with dense forest coverage and shallow snow, that is often found nearby urban areas. In these areas, citizen observations can be fused with satellite-based estimations to deliver more accurate solutions. To that end, we use snow-related tweets that have been annotated by artificial intelligence (AI) methods and are introduced in a novel neural network model, aiming to increase the estimation accuracy of the state-of-the-art remote sensing method. The proposed model combines the estimated snow depth from Sentinel 1 images with the number of Twitter posts and Twitter images that are semantically relevant to snow. The use of instant social media data for purposes of snow depth estimation is investigated, validated and tested in Finland. Our results show that this approach does improve the snow depth estimation, highlighting its potential for use in civil protection agencies in managing snow conditions.

Index Terms—Sentinel 1 images, Backscatter measurements, Snow depth, social media in-situ data, multimodal fusion.

I. INTRODUCTION

NORTHERN European countries, such as Finland, experience a very long winter season which can last several months depending on the region, putting a lot of stress on infrastructure. Also, extreme snow events, besides the benefit it has on winter tourism, can also cause severe problems in electricity supply [1], traffic [2], agriculture and reindeer herding [3]. Addressing these problems requires a well-organized civil protection agency, as well as significant financial resources. In fact, weather and climate related risks, including problems due to snow events are specifically addressed in the Finland's National risk Assessment[1]. Recently, civil protection agencies have adopted the use of remote sensing from a number of platforms (e.g. satellite images and UAVs) in extreme event management, with potential benefits in financial cost and decision making. However, very often the remote sensing data collected by these platforms are not available fast enough for the decision making required in such events. Additionally, malfunction of these platforms can lead to delayed decision

D. Mantsis, M. Bakratsas, S. Andreadis, A. Moumtzidou, I. Gialam-poukidis, S. Vrochidis, and I. Kompatsiaris are with the Centre for Research and Technology Hellas (CERTH), Information Technologies Institute, Thessa-loniki, Greece e-mail: dmantsis@iti.gr, mbakratsas@iti.gr, andreadisst@iti.gr, moumtzid@iti.gr, heliasgj@iti.gr, stefanos@iti.gr, ikom@iti.gr.

P. Karsisto and A. Karppinen are with Space and Earth Observation Centre, Finnish Meteorological Institute, Helsinki, Finland e-mail: pet-teri.karsisto@fmi.fi, Ari.Karppinen@fmi.fi

[1]http://urn.fi/URN:ISBN:978-952-324-249-4

making, very often with tragic consequences. Therefore, complementary data sources may be used to fill this gap. So far a fusion of social media with snow depth detection through remote sensing techniques has not been examined before. An example is given in Figure 1, with publicly available information on Twitter about snow observations, either as textual or visual content.
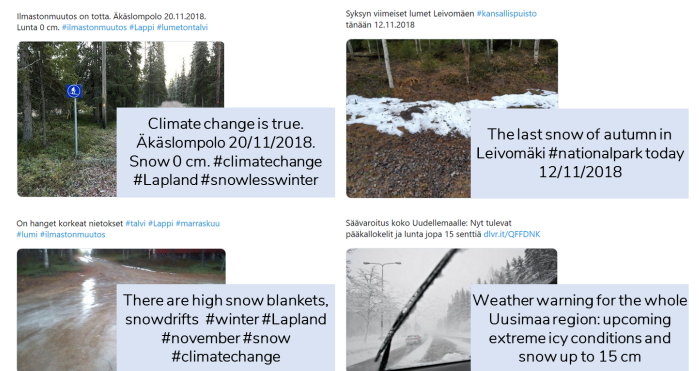

Fig. 1: Example of online available social media data from Twitter posts in Finland

During the last decade, publicly available information from social media (e.g. Twitter, Facebook, Instagram, and various blogs) have made their debut in civil protection agencies around the globe. They can be used as a snap-shot of the public response to an extreme event and they are instant, with a huge benefit to decision making. It turns out that when it comes to breaking news, Twitter appears to have outperformed traditional media. During the Sichuan 2008 earthquake, which was responsible for 70.000 deaths, information regarding an initial tremor was being disseminated by Twitter several minutes before the main burst, and a similar case took place during the 2008 Southern California earthquake as well [4].

Twitter data have been used to disseminate information immediately after the Haiti 2010 earthquake providing an insight regarding injured and trapped victims as well as damaged buildings [5]. Pollution related health issues is another potential field where Twitter could be used given that the public response through this platform is strongly correlated with the ambient air quality in major urban areas [6]. Twitter can also be an effective source of data in other extreme events scenarios. For example, it has been shown as an effective source of data that can be used to identify flood "hot-spots" immediately after these take place, which will, in turn, be used

to task remote sensing data collection (Satellites and UAVs) for a more detailed analysis during the crisis management [7].

Given that microblogging has evolved into such a widespread tool during a crisis, we attempt to expand its use and highlight its potential in managing snowfall events. Our goal is to highlight possible correlation between snow depth and the amount of microblogging associated with this event. This would suggest that the amount of the related information that is shared among the public could be used as a proxy for the intensity of the meteorological event.

Traditionally, SAR C-band backscatter satellite measurements have been used to study snow melt [8], [9], based on the high dielectric losses of water that lead to a reduced backscatter coefficient over wet snow compared to surfaces that are snow-free or covered by dry-snow. For snow depth, C-band satellite backscatter measurements were used early on in the past, but only in co-polarization, and showing only limited sensitivity [10], [11]. On the other hand, cross-polarization backscatter has been used to estimate snow depth in the past, but only locally using tower installations [12]. Recently, a new method that utilizes the ratio of cross-polarization (σ_{vh}^0) to co-polarization backscatter (σ_{vv}^0) has been implemented on Northern Hemisphere mountainous regions, exhibiting a promising snow depth estimate [13]. In our case this methodology is applied in a near-urban area around the city of Helsinki.

Fusing a variety of remote sensing data has been addressed in past studies, for example for land use classification [14], [15], showing that a multi-modal or cross-modal learning provides a significant advantage over single modal applications. In this letter, we use a different approach and we fuse satellite data with social media data, and more specifically we enhance a state-of-the-art snow depth estimation approach [13], with the use of additional in-situ data from citizens (Twitter posts and images). Fusion of remote sensing data with social media is something rather new and will likely draw more attention in the coming years.

II. METHODOLOGY

The proposed methodology (Figure 2) combines Sentinel 1 images and Twitter data that are highly correlated with actual snow conditions.

A. Snow depth estimation using Earth Observation data

a) Pre-processing: Pre-processing of the raw satellite data (Sentinel 1) is carried out through the Sentinel Application Platform (SNAP[2]) and includes thermal noise removal, radiometric calibration, terrain correction and linear to dB transform. Two separate locations are chosen for this in order to validate the results from the satellite derived snow depth (Section III). Given that Southern Finland is covered to a large extent by forests, which attenuate backscatter, we mask out these areas till we are left with the open grassy and farmland areas. Additionally, we mask out water areas (lakes) regardless of whether the water bodies are frozen or not.

In order to properly process the backscatter data that often come from different orbits we need to remove the static bias.

[2]https://step.esa.int/main/toolboxes/snap/

First we average 1-year of backscatter data for each orbit separately, and then average for all orbits. The static bias for each orbit is estimated as the difference between the all orbit mean and the specific orbit mean, which is subtracted from the backscatter time series for the corresponding orbit. Also, before computing their ratio, co-polarized and cross-polarized backscatter are re-sampled and projected onto a coarser $1Km^2$ grid by linear averaging, which also helps reduce speckle noise. If more than 20% of the pixels, corresponding to the native Sentinel 1 resolution, are assigned as water or forest when they are projected onto the coarser grid, the $1Km^2$ average is automatically removed from the analysis.

b) Snow depth estimation from Sentinel 1: The first step in estimating the snow depth is to estimate the snow index [13] as follows:

$$SI(i,t) = \begin{cases} max(0, [SI(i,t-1) + BR(i,t) - BR(i,t-1)]) \\ \qquad\qquad\qquad\qquad\qquad \text{if } SC(i,t) = 1 \\ 0, \qquad\qquad\qquad\qquad\quad \text{if } SC(i,t) = 0 \end{cases}$$

where SC is the snow cover, BR represents the ratio of backscatter radiation in cross-polarization (VH) and co-polarization (VV) if the backscatter is given in linear scale. However, if the backscatter is given in dB the difference must be used instead. We rescale the snow index into snow depth:

$$SD(i,t) = \frac{a}{1 - bFC(i)} SI(i,t) \qquad (1)$$

where $a = 1.1dB^{-1}$ and $b = 0.6$ are constants and FC is the evergreen forest cover fraction (dimensionless) [13].

B. Citizen observations about snow and reporting on social media

1) Representation of Twitter text : Starting with the text representation, we used the state-of-art algorithm Bidirectional Encoder Representation from Transformers (BERT) [16]. BERT involves an attention mechanism that learns contextual relations between words in a text. BERT's goal is to generate a language model (LM), and the used mechanism reads the entire sequence of words at once, contrary to directional models (e.g. n-gram LMs [17], and neural network LMs [18] that read the text input sequentially. Therefore, it is considered non-directional, that allows the model to learn the context of a word based on its surroundings. In order to capture the text representation of the whole tweet, we used an existing pre-trained model in Finnish language called 'bert-base-finnish-cased-v1'. Thus, the input in the BERT model is a Twitter text and the output is a feature vector of length 768.

2) Logistic regression classification in Twitter text (tweet): A Logistic Regression (LR) model is trained to classify Twitter posts (short text) as relevant or not to snow. This disambiguation allows the removal of tweets that refer to metaphoric meanings of the word "snow" or synonyms of it. The model is trained by using the manually annotated data provided by domain experts (i.e. Finnish Meteorological Institute) and a grid search is realized in order to identify the best parameters. We report that a basic logistic regression model is developed with the inverse of regularization strength parameter, usually
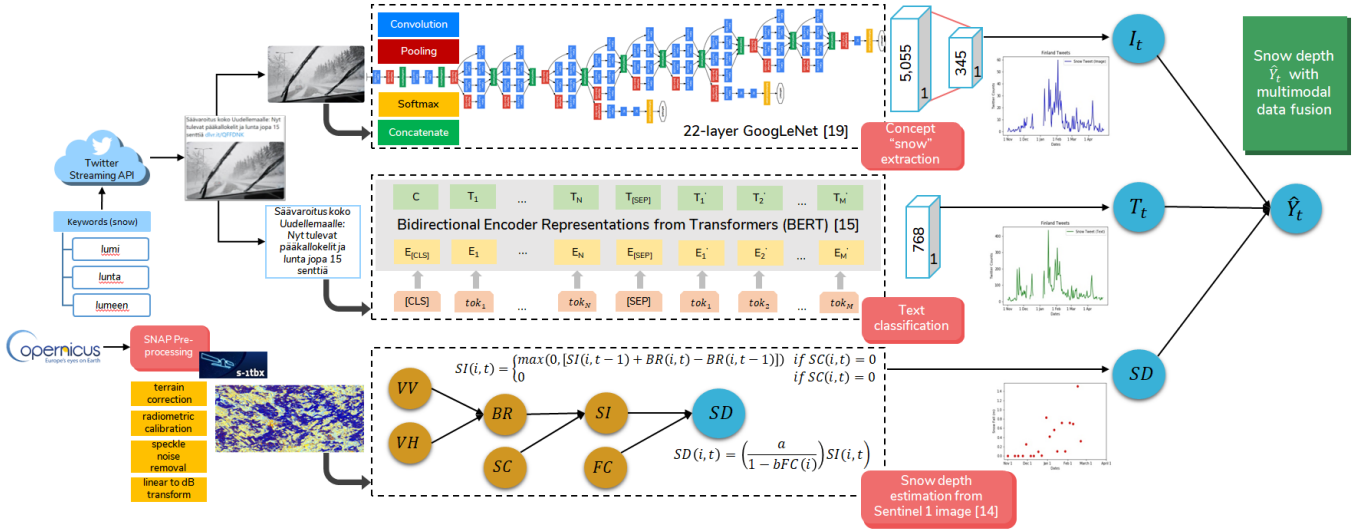
Fig. 2: The proposed framework for snow depth estimation using Earth Observation (EO) and social media data. The visual concept *snow* is extracted, the relevance of the Twitter text to the concept *snow* is estimated from text classification and the snow depth is estimated from Sentinel 1 images.

denoted by C, equal to 31.57. This model was trained and validated on a dataset of 7,710 tweets, out of which 3,877 where annotated as irrelevant (50.3%).

3) Deep Learning for visual concept (snow) extraction on Twitter images: A framework different to text classification of the tweets has been used to extract visual concepts from the subset of tweets that include images. The target is to count the number of Twitter images per day that illustrate snow. To that end, we used a 22-layer GoogleNet network [19] that was trained on 5,055 ImageNet concepts [20]. At this stage, the classification layer of the network, which is a fully connected layer, has dimension equal to 5,055, i.e. the total number of given concepts. Additional fine-tuning is performed so as to reduce the number of concepts to the 345 SIN TRECVID concepts[3], which involves replacing the classification layer with dimensionality 5055 with a fully connected classification layer with dimension equal to 345. The concept "snow", which is defined as: *Snow falling or already accumulated on the ground*) is one of the TRECVID 345 concepts. Thus, this framework receives as input an image, then the fine-tuned Deep Convolutional Neural Network (DCNN) is tested on the specific image and a list of concepts along with their probabilities is produced. This framework, was trained on two datasets: a) the TRECVID SIN 2013 dataset and b) the PASCAL VOC 2012 dataset, achieving a Mean Average Precision of 80%. If the concept "snow" is ranked among the top-10 concepts with probability higher than 0.01, then we consider that the image contains the concept. Such values are very common in concept extraction, where our neural network obtain probability values even less than 10^{-4} in case a concept does not exist in the image. Concepts are extracted from all the Twitter images, and 1,118 have been found to contain the concept "snow".

C. Correlating social media observations with snow depth

The snow depth estimate can be, in some cases, strongly correlated with the actual measurements, and weakly correlated in other. Even in the case when the correlation is significant, severe biases may exist due to an overestimation or underestimation of the actual measurements. To reduce these biases we use regression analysis to compute a new estimate of the snow depth, $\hat{Y}_t$:

$$\hat{Y}_t = \alpha_1 SD(t) + \alpha_2 I_t + \alpha_3 T_t + \beta \quad (2)$$

where SD is the estimated snow depth from Sentinel 1 images only [13]. Moreover, we denote by T_t and I_t the number of relevant tweets per day t, or the number of Twitter images that contain the concept snow, respectively. The coefficients α_1, α_2, α_3, and β are constants.

The proposed model of Eq. (2) is using $SD(t)$, T_t and I_t, showing the added value of social-generated data assets, which is either estimated from Logistic regression classification in Twitter text or a DCNN model on Twitter images to count the number of posts that are relevant to snow.

III. EXPERIMENT AND RESULTS

A. Dataset description

Regarding snow depth, we use in-situ measurements from four sites around the city of Helsinki (Figure 3), for validation purposes. All four sites are located in areas with intense urban development, which is ideal for this study given that we need to accurately represent the snowfall conditions of the area where the public response (Twitter) takes place.

11,024 tweets were collected, covering a period of 151 days, i.e. from November 2018 till March 2019. The Twitter Streaming API[4] is used to collect relevant Twitter posts. The

---

[3]https://www-nlpir.nist.gov/projects/tv2015/index.html#sin

[4]https://developer.twitter.com/en/docs/tweets/filter-realtime/overview

Fig. 3: Area of interest and measurements from four pilot sites in Finland.

collection is keyword-based, where Finnish words for snow (e.g., lumi, lunta, lumeen) have been used as queries. This dataset was manually annotated and was split into a training and testing dataset consisting of 5783 and 1927 respectively. The mean precision, recall, accuracy and F-score that were achieved are 73,28%, 73,58%, 73,5% and 73,43% respectively. Eventually, the model is validated on a set of 11,024 new non-annotated data, and the model estimated 6,097 tweets (55.3%) as relevant and 4,927 (44.7%) as irrelevant.

B. Correlation between social media and snow depth

From a close inspection of the daily in situ snow depth measurements (not shown) and the Twitter time series (text-based or image-based) we can see that the three largest snow events (during 09, 17, and 29 January 2019) coincide with peaks in tweeted snow images and number of relevant tweets which took place the same or the next day. We denote by $\rho(A, B)$ the correlation coefficient between two variables A and B. The corresponding estimated Pearson correlation is denoted by $r(A, B)$, which results the following outcomes with the average in situ snow depth Y:

- $r(Y, I_t) = 0.56$, positive and statistically significant
- $r(Y, T_t) = 0.51$, positive and statistically significant

C. Data fusion between social media and Sentinel 1 images

The satellite derived snow depth of Eq. (1) is estimated only based on Sentinel 1A measurements during Dec 2018-March 2019. During the study period (2019 winter), Sentinel 1B satellite did not give any measurements for the area of Helsinki in IW mode, and are given in EW mode instead due to the Baltic Sea Ice campaign, limiting our snow depth estimate to every 6 days only in area A and with a slightly increased frequency in area B. Additionally, we exclude the snow depth computation during the period between late February to March, given the fact that the liquid water from melting snow during the melting period overwhelms any backscatter signal from shallow snow. The resulting estimated snow depth is poorly correlated with the in situ measurements, i.e. $r = 0.13$ for area A and $r = 0.41$ for area B. This is most likely due to the fact that the snow
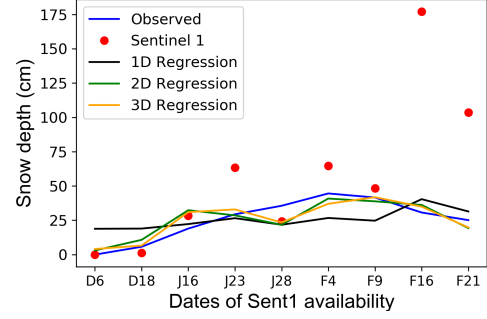
(a) Area A
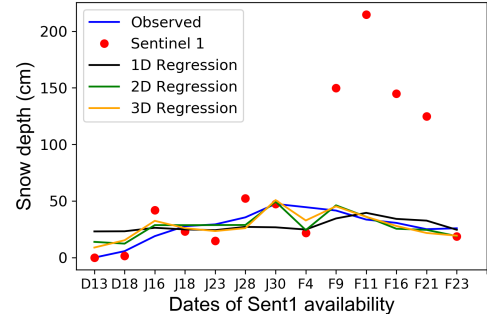
(b) Area B

Fig. 4: Variation of in situ and satellite derived snow depth. The estimated snow depth based on the 1D, 2D and 3D regression model is also shown. The data are only presented for the dates when the Sentinel 1 backscatter is available.

was shallow in the area of Helsinki, which did not allow the [13] method to exhibit its full potential. However, here is exactly where the use of Twitter data are fused with the satellite-based estimation to provide an improved estimation. The improvement is measured with the Mean Squared Error (MSE), that represents the deviation between the in situ snow depth measurements with each model.

The positive correlation between the tweets and the snow observations is a necessary condition that allows us to fuse these two through a regression model. This can be seen in more detail as we build the regression model, based on the correlation between the in situ snow depth and the number of tweets (variables I_t and T_t).

The Sentinel 1 derived snow depth exhibits a large mean square error (MSE) when compared to the actual snow depth. This is mainly the result of outlier points like the one during February the 16th (Figure 4). A 1st order (1D) regression model, which uses only the estimated snow depth, can remove the effect of such outliers by utilizing the linear relationship between the predictor (Sentinel-1 derived snow depth) and the in situ snow depth Y. The resulting snow depth $\hat{Y}$ has a much smaller MSE with no outliers. However, what needs to be examined is whether this can be improved by adding the information given by social media. It turns out that a 2nd order model (2D) expressed by the equation

$$\hat{Y}_t = 0.17SD(t) + 1.27I_t + 0.83 \qquad (3)$$

exhibits a smaller MSE for area A, reduced by 67% compared to the 1st order model, highlighting the importance of social data. The regression coefficients show that the added value is mainly attributed to the Twitter data, which is because this variable has a larger correlation with the in situ snow depth. If we use the text-based tweets instead, the reduction in MSE for the same area is similar, but with a different regression equation:

$$\hat{Y}_t = 0.15SD(t) + 0.20T_t - 1.47 \tag{4}$$

The question raised here is whether using both Twitter variables T_t and I_t at the same time would make any sense. The resulting regression model is:

$$\hat{Y}_t = 0.16SD(t) + 0.60I_t + 0.13T_t - 2.32 \tag{5}$$

which further reduces the MSE by 10% for area A, as it is shown in Table I. This indicates that these variables can be used together. We can also see that the contribution of the information from social media is not shared equally, with a much larger contribution from the image-based tweets. This is due to the fact that I_t has a larger correlation with the actual snow, and a smaller weight is given to T_t which only accounts for the added skill to I_t.

The comparison of the models, presented in Equations (3), (4) and (5) with respect to MSE are presented in Table I.

Snow depth $\hat{Y}(cm)$	MSE (Area A)	MSE (Area B)
$\hat{Y}_t = SD(t)$	3265.85	5314.46
$\hat{Y}_t = \alpha_1 SD(t) + \beta$	164.23	152.68
$\hat{Y}_t = \alpha_1 SD(t) + \alpha_2 T_t + \beta$	60.81	68.81
$\hat{Y}_t = \alpha_1 SD(t) + \alpha_2 I_t + \beta$	54.64	54.99
$\hat{Y}_t = \alpha_1 SD(t) + \alpha_2 I_t + \alpha_3 T_t + \beta$	**47.11**	**54.98**

TABLE I: Mean square error (MSE) between in situ snow depth and the proposed snow depth estimates.

IV. CONCLUSION

The goal of this letter is to evaluate the potential for using satellite backscatter measurements to estimate snow depth in areas that experience a shallower snow and are much more densely forested, and how we can combine these with instant social media data to augment our estimate. Citizens act as sensors generating multimodal in-situ data with short text and informative images that are further utilized to enhance state-of-the-art models for snow depth estimation. Results show that despite the limitations of the backscatter methodology to estimate snow depth, significant improvement can be achieved through the use of regression analysis and social media data. At first, the linear relationship between the snow depth estimate and the in situ snow depth, allows the removal of most of the bias and the noise. To a second degree a fusion of Twitter data into the regression model allows an additional bias removal, which is the epicenter of this work.

ACKNOWLEDGMENT

REFERENCES

[1] R. Molarius, M. Räikkönen, K. Forssén, and K. Mäki, "Enhancing the resilience of electricity networks by multi-stakeholder risk assessment: The case study of adverse winter weather in finland," *Journal of Extreme Events*, vol. 3, no. 04, p. 1650016, 2016.

[2] I. Juga, M. Hippi, D. Moisseev, and E. Saltikoff, "Analysis of weather factors responsible for the traffic 'black day' in helsinki, finland, on 17 march 2005," *Meteorological applications*, vol. 19, no. 1, pp. 1–9, 2012.

[3] M. T. Turunen, S. Rasmus, M. Bavay, K. Ruosteenoja, and J. Heiskanen, "Coping with difficult weather and snow conditions: Reindeer herders' views on climate change impacts and coping strategies," *Climate Risk Management*, vol. 11, pp. 15–36, 2016.

[4] J. Li and H. R. Rao, "Twitter as a rapid response news service: An exploration in the context of the 2008 china earthquake," *The Electronic Journal of Information Systems in Developing Countries*, vol. 42, no. 1, pp. 1–22, 2010.

[5] O. Oh, K. H. Kwon, and H. R. Rao, "An exploration of social media in extreme events: Rumor theory and twitter during the haiti earthquake 2010." in *Icis*, vol. 231, 2010, pp. 7332–7336.

[6] S. Gurajala, S. Dhaniyala, and J. N. Matthews, "Understanding public response to air quality using tweet analysis," *Social Media+ Society*, vol. 5, no. 3, p. 2056305119867656, 2019.

[7] G. Cervone, E. Sava, Q. Huang, E. Schnebele, J. Harrison, and N. Waters, "Using twitter for tasking remote-sensing data collection and damage assessment: 2013 boulder flood case study," *International Journal of Remote Sensing*, vol. 37, no. 1, pp. 100–124, 2016.

[8] T. Nagler, H. Rott, E. Ripper, G. Bippus, and M. Hetzenecker, "Advancements for snowmelt monitoring by means of sentinel-1 sar," *Remote Sensing*, vol. 8, no. 4, p. 348, 2016.

[9] T. Nagler, H. Rott, G. Schwaizer, J. Ossowska, J. Nemec, and U. Fasching, "Operational monitoring of alpine snow cover within the european copernicus programme," in *Proceedings of the International Snow Science Workshop*, vol. 1, no. 1, 2018, pp. 348–252.

[10] M. Bernier, J.-P. Fortin, Y. Gauthier, R. Gauthier, R. Roy, and P. Vincent, "Determination of snow water equivalent using radarsat sar data in eastern canada," *Hydrol. Proces.*, vol. 13, no. 18, pp. 3041–3051, 1999.

[11] J. Shi and J. Dozier, "Estimation of snow water equivalence using sir-c/x-sar. ii. inferring snow depth and particle size," *IEEE Transactions on Geoscience and Remote sensing*, vol. 38, no. 6, pp. 2475–2488, 2000.

[12] J. R. Kendra, K. Sarabandi, and F. T. Ulaby, "Radar measurements of snow: Experiment and analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 864–879, 1998.

[13] H. Lievens, M. Demuzere, H.-P. Marshall, R. H. Reichle, L. Brucker, I. Brangers, P. de Rosnay, M. Dumont, M. Girotto, W. W. Immerzeel *et al.*, "Snow depth variability in the northern hemisphere mountains observed from space," *Nature communications*, vol. 10, no. 1, pp. 1–12, 2019.

[14] D. Hong, N. Yokoya, G.-S. Xia, J. Chanussot, and X. X. Zhu, "X-modalnet: A semi-supervised deep cross-modal network for classification of remote sensing data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 167, pp. 12–23, 2020.

[15] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, Q. Du, and B. Zhang, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[17] R. Rosenfeld, "Two decades of statistical language modeling: Where do we go from here?" *Proceedings of the IEEE*, vol. 88, no. 8, pp. 1270–1278, 2000.

[18] T. Mikolov, M. Karafiát, L. Burget, J. Černockỳ, and S. Khudanpur, "Recurrent neural network based language model," in *Eleventh annual conference of the international speech communication association*, 2010.

[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[20] N. Pittaras, F. Markatopoulou, V. Mezaris, and I. Patras, "Comparison of fine-tuning and extension strategies for deep convolutional neural networks," in *International Conference on Multimedia Modeling*. Springer, 2017, pp. 102–114.