Augmentation

“Bake translational invariances into the dataset such that the resulting models will perform well.”

“Augmentation minimizes the distance between training and validation set, as well as any future dataset.”

“Augmentation can prevent from learning irrelevant patterns / biases.”

Background

(disadvantage of CNN: require big data (23, 24 in Shorten) such that no overfitting (more in This leads to…). High resolution data collection representative for the entire Arctic challenging and labelling pond imagery expensive. Different regularization techniques).

Augmentation refers to synthetically modifying data before model training. By applying image transformations, dataset size can be increased and more diversity added to the training data (Shorten et al). With augmentation methods such as geometric or brightness changes, varying conditions can be simulated that are not captured by the flight at hand, accounting for variations in flight height, atmospheric effects or seasonal and region-based changes, thus improving the generalization of the model especially for future work.

Transformations can happen on pixel-level (such as brightness change or noise injection (57 and 58 in Shorten) or applied on the whole image (flip, rotate, shift) and also targeted towards masks.

(Divide augmentation methods and list some: Basic augmentation techniques are geometric transformations, color space transformations and kernel filters).

(explicitly state that augmentation is applied on training data only).

Augmentation can be applied at to stages in model training: ‘Offline’ before training (increases dataset size, requires memory and training time, preferred for smaller datasets (source)) or ‘on the fly’, generating augmented samples during each training epoch, just before feeding into the model.

Augmentation has been proven as one of the most powerful techniques to combat overfitting, especially in cases with small data. It has been successfully applied in Remote Sensing tasks, an overview can be found in (reference).

Methodology

Data augmentation has been considered to increase the dataset size and variety of training information. However, inappropriate augmentation can introduce unrealistic transformations to the dataset and result in decreasing model performance. To test the effectiveness of different techniques, a pool of preselected methods has been created. The number of augmentation methods applied has been increased incrementally. If a method resulted in decreasing model performance, it was disregarded for further configurations. If it didn’t lead to improvement nor decrease, it was kept to account for more variety in future data. To safe training time, methods have been tested in ‘on fly’ mode. Augmentations have been implemented using Albumentations library.

The methods considered were the following, added incrementally:

* Rotation, horizontal and vertical flipping: Simulate changes of orientation. As the images are in overhead perspective, labels are preserved.
* Cropping: Simulate instabilities in flight height and introduce variations in object position and scale.
* Brightness and Contrast: Brightness changes simulate varying surface temperatures while contrast could capture seasonal changes, when temperature differences across ice and water get greater.
* Sharpening and blurring: Blurring (and motion blurring) simulates effects of noise or atmospheric influences such as water vapour. Sharpening could reduce these effects. Sharpening and blurring have been applied exclusively to one another.
* Noise Injection: Gaussian noise has been added to increase robustness to noise.

Each method in the current configuration has been applied with a probability of 0.5 to each training image. For methods that apply interpolation, interpolation has been changed to nearest to preserve categorical mask labels. Other parameters were kept default. Note that for rotation, reflection is used as border interpolation / background embedding method which was decided to be the most natural but however might result in artificial elongations of floe and pond features at the image border.

Color augmentations such as jittering have been disregarded to preserve labels. RGB-based methods were not applicable due to one-channel. Perspective changing transformations have been disregarded as images will be usually taken of the same angle (overhead perpective).

To test the effectiveness of on-fly versus offline augmentation, the best methods have been selected and applied for both trainings. On-fly augmentation is directly applied when retrieving data and will result in more efficient model training, however, no additional datasize increase. In offline augmentation, one can choose the magnitude by which dataset size is increased. As there is no consensus as to which ratio of original and augmented data is best (Shorten), three different factors have been tested: 2, 5 and 20. Note that a larger factor results in more memory requirements and longer training time for the same number of epochs.

Results

* Rotation, cropping, flipping: Although regarded as most effective in literature (63 and 119 in Shorten), no increase in performance could be observed. Rotation performs best: 19 and 23 in Shorten.
* Brightness Contrast resulted in unstable training and worse performance and was therefore disregarded for this study. However, for future applications, with regard to a more varying application domain, it could be an option. The same accounts for Gaussian noise and sharpening and blurring.
* Also especially cropping could be helpful for future datasets, when flight height is not constant, and blurring sharpening for different spatial resolutions.

Discussion

* For the scope of this thesis, investigations have been reduced to single-image augmentation tools. More involved methods like mixing images (Mixup, Cutmix) or deep learning methods (feature space augmentations, adversarial training, neural style transfer, GANs, meta learning data augmentation) that have been shown successful in other work (references). However, note label preservations (cutmix). Random Erasing. To get a more robust prediction, test-time augmentation could be used (120, 121 in Shorten).
* Shearing transformation could be applied to account for distortions, same: Elastic deformations.

FIGURES

Figure that shows augmentation example.

Results.

TO-DO:

* Wie genau wird on-fly augmentierung angewendet?