

	ブランド ザーラ
氏 名	BRANDT SARAH
本籍（国籍）	ドイツ連邦共和国
学位の種類	博士（農学）
学位記番号	連研第 804 号
学位授与年月日	令和 3 年 9 月 2 4 日
学位授与の要件	学位規則第 5 条第 1 項該当課程博士
研究科及び専攻	連合農学研究科 地域環境創生学専攻
学位論文題目	Plant species detection and classification with image analysis focused on deep learning techniques in natural environments （深層学習技術を用いた画像分析による自然環境下における樹種の識別と分類）
学位審査委員	主査 山形大学教授 Lopez Caceres Maximo Larry 副査 山形大学准教授 吉村 謙一 副査 弘前大学助教 鄒 青穎 副査 岩手大学准教授 武藤 由子

論 文 の 内 容 の 要 旨

Forests, with an area of 4.15 billion hectares, cover large areas of the earth surface and provide essential economic, environmental, and social services, like balancing carbon cycles, conserve biodiversity and regulate fresh water supplies. Human disturbances and climate change affect forest ecosystems, their growth and resilience. Changing forest stresses and changes in the ecosystem increased the demand of accurate forest information. This is especially important in Japan, due to the sheer magnitude and complexity of the forest area, composed of natural mixed broadleaved deciduous forests.

Traditionally, information about forests has been collected using expensive and work-intensive field inventories collecting data about tree species compositions, distributions and forest structures by identifying trees and counting them. These and further information are needed to do sustainable forest management and monitoring of the ecosystems. Challenging environments restricted forest inventories to be conducted in small field plots and data were scaled-up to forest stand levels. In natural mixed forests small field plots were found to be less representative for the forests studied.

In recent years Unmanned Aerial Vehicles (UAVs) have become very popular as they represent a simple and inexpensive way to do automated flid surveys by gathering high resolution data of large forested areas. Among the many sensors that UAVs can carry, RGB cameras are fast, cost-effective and make the data gathering and interpretation easy, as they do not need intensive calibrations or pre-processing. With spatial resolution of less than 1 m, canopy structures and even leave structures can be identified, so that tree species classification can be performed from images. In addition, deep learning (DL) has also been catching more attention in the field of forestry. In particular as a way to include the knowledge of

forestry experts into automatic software pipelines, to deal with tree detection, tree health or tree species classification. The increasing numbers of studies performed in forestry using DL together with improving software and data availability increased the need of such kind of studies in Japan's forests.

In this study three methodologies were applied to characterise and evaluate natural mixed forests in Japan regarding their tree species composition, distribution and structure: Field surveys, Image analysis with GIS application and deep learning. These three methodologies were used for the main analysis of the Yamagata University Research Forest (YURF) sites, as one aspect was that the deep learning results needed to be evaluated regarding their accuracy compared with field inventories. The other aspect was that the collected images could further be used for image analysis with Geographic Information Systems (GIS). Therefore, field data were collected in the traditional way of classifying tree species and counting them in the field. Data were collected for seven sites with distribution and survey maps, while for three slope sites only tree counting information were provided. For the total of 13 sites, image analyses were performed using manual annotations and GIS applications generating tree survey, tree distribution, density, count, hot spot, DEMs, aspect and slope maps for the different sites and tree species. In total, 70 tree species were identified in the field, while 41 species classes were identified from the images. Field and image results were analysed regarding biodiversity, densities, frequencies and dominant tree species in the different considered sites. Results of the image analyses and the field analyses were similar, when canopy and subcanopy species were considered; differences were mainly found for understory vegetation and shrubs, which were hardly identifiable from images. Significant differences occurred only in the slope sites, where small *Quercus mongolica* and *Fagus crenata* trees were covered by canopy layer species, but still dominated those forests, according to the field data. The image data identified several species as dominant, depending on the observed site (*Magnolia obovata*, small-leaved *Acer*, *Acer mono maxim* and *Quercus mongolica*). In riparian and terrace sites, *Juglans ailantifolia*, *Pterocarya rhoifolia*, *Salix* species and *Acer* species dominated the sites. Biodiversity measures indicated a higher diversity when field data were used, in contrast to image data. While the Shannon diversity values ranged between 1.73 and 2.39 (with the evenness ranging between 55 % and 82 %) for the image analysis results, it ranged between 2.14 and 2.76 (evenness: 71 % - 84 %). Layering of the forest was better classified with the field data, as all layers could be easily identified, while most of the lower vegetation could not be identified from images. The forest sites were classified based on the layering and the dominance of tree species, while the dominance of canopy areas was used for the classification. Riparian and terrace sites were *Juglans ailantifolia* forests with *Pterocarya rhoifolia* and *Salix* as co-dominant species. According to the results of the image analyses, *Juglans ailantifolia*, *Acer* species, *Magnolia obovata* and *Quercus mongolica* were the dominant species, which confirmed the field analysis, but with a different order (small-leaved *Acer* and *Quercus mongolica* were the majority). The evaluation of the images with GIS tools enhanced the visibility of important aspects

hidden in the data, and spatial information could be easily extracted and interpreted.

Generally, image analyses of tree canopy areas provided more accurate information than tree counting, as dense canopy areas hampered counting. To evaluate semi-automatic counting, field data, DEMs, summer and winter images were used to count tree species. Highest accuracies were reached when counting was performed on winter images, while small and young trees still remained challenging to count.

The main aim was to classify plant species automatically with deep learning techniques. Therefore, data were gathered in the 13 sites in the YURF in four seasons and over three years, as well as in the coastal forest near Sakata city and in Lichtenmoor, a wetland area in Germany. The 13 sites in YURF were divided into Riparian, Terrace and Slope sites. All images were processed with Metashape to produce orthomosaics and DEMs (Digital Elevation Models). In a first step, deep learning was applied to a simple example of classifying trees with leaves versus trees without leaves (deciduous vs. evergreen) in winter images, to assess the effect of transfer learning and deep learning architectures (ResNet50 and UNet). In this approach also multi-label patch (MLP) classification versus semantic segmentation were studied, breaking the orthomosaics into image patches. The results showed that transfer learning is necessary to achieve satisfactory outcome with MLP classification of deciduous versus evergreen trees. In the winter orthomosaic dataset the improvement from no transfer learning to transfer learning from a general-purpose dataset was 9.78 %. Furthermore, the ResNet50 architecture showed a high performance with better results than the UNet. The results indicated already that data balancing is an important topic. The study of invasive blueberry species, endangering sensitive wetland environments and black locusts invading into coastal forests, were two more application-oriented examples, with an easier problem definition: Classifying target species in a natural green environment. The ResNet50 architecture was used with transfer learning to detect black locust trees in an evergreen coniferous black pine forest with a 75 % of True Positives (TP) and 9 % False Positives (FP) while the detection of native trees achieved 95 % TP and 10 % FP. Detections of invasive blueberry bushes were performed with ResNet50, transfer learning and unfrozen weights with True Positive Values (TPV) of 93.83 % and an Overall Accuracy (OA) of 98.83 %. A refinement of the result masks reached a Dice of 0.624. Image analyses were performed to produce maps of blueberry location, distribution and spread in each study site, as well as bush height and area information. A preliminary study of different deep learning networks, transfer learning, the use of data augmentation and loss functions and settings were tested for the detection of invasive blueberry species. The challenge of the data was the imbalance, as invasive species had fewer individuals than natural occurring plants. In this study of state-of-the-art deep learning architectures the best results were obtained with the ResNeXt architecture (93.75 True Positive rate), and 98.11 % accuracy for the Blueberry class with ResNet50; Densenet and wideResNet calculated similar results. The knowledge, gained with easy examples, was then applied to automatic

tree species classification in natural mixed forest. This study provides an efficient and effective methodology to study forests and other natural environments, like wetlands, using different techniques: field surveys, image analyses and deep learning. Automatic generated results showed high accuracies and indicated the applicability of the methodology in different fields. Image analyses extract the most important information of aerial images, depending on the study focus. Field data captured a lot of information that could not be extracted from images, and therefore the methodology set provided new and important insights into forest environments.

(日本語)

地球上で 41 億 5,000 万ヘクタールの森林は、地表の広い範囲を占め、炭素循環の調整、生物多様性を保全、水供給など、経済的、環境的、社会的に不可欠なサービスを提供しています。人為的な攪乱や気候変動は、森林生態系における成長や回復力に影響を与えます。従来、森林に関する情報は、コストと労力のかかる毎木調査によって収集されてきました。毎木調査は通常小規模なフィールドプロットで行われ、そのデータを用いて森林レベルにスケールアップされてきました。しかし、日本に多く分布している天然生広葉樹林は非常に複雑であるため、小さなフィールドプロットでは調査対象の森林をあまり代表していないという問題が生じていました。

近年、UAV (Unmanned Aerial Vehicle : 無人航空機) が森林調査において用いられるようになってきました。UAV が搭載できる RGB カメラは集中的なキャリブレーションや前処理を必要としないため、データの収集と解釈が容易です。また、1m 以下の空間分解能で、林冠構造や葉の構造まで識別できるため、画像から樹種分類を行うことが可能です。

本研究では、現地調査、GIS を用いた画像解析、ディープラーニングという 3 つの方法論を用いて、日本の天然生混合林の樹種構成、分布、構造を特徴づけ、評価した。山形大学研究林 (YURF) の合計 13 サイトにおいて、フィールドデータを取得し、GIS アプリケーションを用いて画像解析を行いました。その結果、合計でフィールドデータからは 70 種の樹木が確認され、画像データを用いた深層学習からは 41 種のクラスが特定されました。林冠種を考慮した場合、画像分析とフィールド分析の結果は類似していましたが、下層植生と低木においては両結果に差異が生じ、画像からはほとんど識別できませんでした。生物多様性の指標は、画像データとは対照的に、フィールドデータを使用した場合に高い多様度を示した。画像解析の結果では、シャノン多様性指数が 1.73~2.39 (均等性は 55%~82%) であったのに対し、2.14~2.76 (均等性は 71%~84%) であった。森林の階層構造に関して、画像からは下層においては識別できませんでした。溪畔林に設置したサイトは、実際にはオニグルミが優占する森林で、サワグルミが共優占種であったが、画像分析の結果によると、オニグルミ、カエデ属、ホオノキ、ミズナラが優占種であると認識され、優占度の順序が異なっていました。GIS ツールを使って画像を評価することで、データに隠れていた重要な側面が見えやすくなり、空間情報を簡単に抽出して解釈することができました。

一般的に、樹冠が密集していると個体数推定が困難になることが知られています。そこで、フィールドデータ、DEM、夏季および冬季の画像を用いて樹木個体数の正確な評価を半自動解析によりおこないました。その結果、冬の画像でカウントを行った場合に最も高い精度が得られましたが、小さい木や若い木のカウントは依然として困難でした。

次に深層学習技術を用いて植物種を自動的に分類するために、YURF の 13 のサイト、酒田市近郊の海岸林や、ドイツの湿地帯でもデータを収集しました。すべての画像は Metashape で処理され、オルソモザイクと DEM (Digital Elevation Models) を作成しました。まず、冬の画像で葉のある木とない木 (落葉樹と常緑樹) を分類するという簡単な例に深層学習を適用し、伝達学習と深層学習のアーキテクチャ (ResNet50 と UNet) の効果を評価しました。このアプローチでは、マルチラベルパッチ (MLP) 分類とセマティックセグメンテーションについても検討しました。その結果、MLP による落葉樹と常緑樹の分類で満足いく結果を得るためには、転移学習が必要であることがわかりました。冬のオルソモザイクのデータセットでは、転移学習を行わなかった場合と、汎用データセットからの転移学習を行った場合の改善率は 9.78% でした。さらに、ResNet50 アーキテクチャは、UNet よりも優れた結果を示し、高い性能を発揮しました。ResNet50 アーキテクチャを伝達学習とともに使用して、常緑針葉樹のクロマツ林におけるニセアカシアの木を検出したところ、真陽性 (TP) が 75%、偽陽性 (FP) が 9% であったのに対し、在来種の木の検出では TP が 95%、FP が 10% となりました。

本研究では、森林や湿地帯などの自然環境を調査するための効率的かつ効果的な方法論を、現地調査、画像解析、深層学習などのさまざまな手法を用いて提供することができました。自動生成された結果は高い精度を示し、さまざまな分野での方法論の適用可能性を示しました。画像解析では、調査対象に応じて航空画像の中から最も重要な情報を抽出します。フィールドデータは、画像からは抽出できない多くの情報を捉えているため、この方法論セットは、森林環境に関する新たな重要な洞察を提供するものだと考えられます。

論文審査の結果の要旨

Forests, with an area of 4.15 billion hectares, cover large areas of the earth surface and provide essential economic, environmental, and social services, like balancing carbon cycles, conserve biodiversity and regulate fresh water supplies. Human disturbances and climate change affect forest ecosystems, their growth and resilience. Changing forest stresses and changes in the ecosystem increased the demand of accurate forest information. This is especially important in Japan, due to the sheer magnitude and complexity of the forest area, composed of natural mixed broadleaved deciduous forests.

Traditionally, information about forests has been collected using expensive and work-intensive field inventories collecting data about tree species compositions, distributions, and forest structures by identifying trees and counting them. These and further information are needed to do sustainable forest management and monitoring of the ecosystems. Challenging environments restricted forest inventories to be conducted in small field plots and data were scaled-up to forest stand levels. In natural mixed forests small field plots were found to be less representative for the forests studied.

In recent years Unmanned Aerial Vehicles (UAVs) have become very popular as they represent a simple and inexpensive way to do automated flid surveys by gathering high resolution data of large forested areas. Among the many sensors that UAVs can carry, RGB cameras are fast, cost-effective and make the data gathering and interpretation easy, as they do not need intensive calibrations or pre-processing. With spatial resolution of less than 1 m, canopy structures and even leave structures can be

identified, so that tree species classification can be performed from images. In addition, deep learning (DL) has also been catching more attention in the field of forestry. In particular, the inclusion of knowledge from forestry experts into automatic software pipelines to deal with tree detection, tree health or tree species classification. The increasing numbers of studies performed in forestry using DL together with improving software and data availability increased the need of such kind of studies in Japan's forests.

In this study three methodologies were applied to characterize and evaluate natural mixed forests in Japan regarding their tree species composition, distribution, and structure: Field surveys, Image analysis with GIS application and deep learning. These three methodologies were used for the main analysis of the Yamagata University Research Forest (YURF) sites, as one aspect was that the deep learning results needed to be evaluated regarding their accuracy compared with field inventories. The other aspect was that the collected images could further be used for image analysis with Geographic Information Systems (GIS). Therefore, field data were collected in the traditional way of classifying tree species and counting them in the field. Data were collected for seven sites with distribution and survey maps, while for three slope sites only tree counting information were provided. For the total of 13 sites, image analyses were performed using manual annotations and GIS applications generating tree survey, tree distribution, density, count, hot spot, DEMs, aspect and slope maps for the different sites and tree species. In total, 70 tree species were identified in the field, while 41 species classes were identified from the images. Field and image results were analyzed regarding biodiversity, densities, frequencies, and dominant tree species in the different considered sites. Results of the image analyses and field analyses were similar, when canopy and sub-canopy species were considered; differences were mainly found for understory vegetation and shrubs, which were hardly identifiable from images. Significant differences occurred only in the slope sites, where small *Quercus mongolica* and *Fagus crenata* trees were covered by canopy layer species, but still dominated those forests, according to the field data. The image data identified several species as dominant, depending on the observed site (*Magnolia obovata*, small-leaved *Acer*, *Acer monomaxim* and *Quercus mongolica*). In riparian and terrace sites, *Juglans ailantifolia*, *Pterocarya rhoifolia*, *Salix* species and *Acer* species dominated the sites. Biodiversity measures indicated a higher diversity when field data were used in contrast to image data. While the Shannon diversity values ranged between 1.73 and 2.39 (with the evenness ranging between 55% and 82%) for the image analysis results, it ranged between 2.14 and 2.76 (evenness: 71% -84%). Layering of the forest was better classified with the field data, as all layers could be easily identified, while most of the lower vegetation could not be identified from images. The forest sites were classified based on the layering and the dominance of tree species, while the dominance of canopy areas was used for the classification. Riparian and terrace sites were *Juglans ailantifolia* forests with *Pterocarya rhoifolia* and *Salix* as co-dominant species. According to the results of the image analyses, *Juglans ailantifolia*, *Acer* species, *Magnolia obovate* and *Quercus mongolica* were the dominant species, which confirmed the field analysis, but with a different order (small-leaved *Acer* and *Quercus mongolica* were the majority). The evaluation of the images with GIS enhanced the visibility of important aspects hidden in the data, and spatial information could be easily extracted and interpreted.

Generally, image analyses of tree canopy areas provided more accurate information than tree counting, as dense canopy areas hampered counting. To evaluate semi-automatic counting, field data, DEMs,

summer and winter images were used to count tree species. Highest accuracies were reached when counting was performed on winter images, while small and young trees still were difficult to count. The main aim was to classify plant species automatically with deep learning techniques. Therefore, data were gathered in the 13 sites in the YURF in four seasons and over three years, as well as in the coastal forest near Sakata city and in Lichten moor, a wetland area in Germany. The 13 sites in YURF were divided into Riparian, Terrace and Slope sites. All images were processed with Metashape to produce orthomosaics and DEMs (Digital Elevation Models). In a first step, deep learning was applied to a simple example of classifying trees with leaves versus trees without leaves (deciduous vs. evergreen) in winter images, to assess the effect of transfer learning and deep learning architectures (ResNet50 and UNet). In this approach also multi-label patch (MLP) classification versus semantic segmentation were studied, breaking the orthomosaics into image patches. The results showed that transfer learning is necessary to achieve satisfactory outcome with MLP classification of deciduous versus evergreen trees. In the winter orthomosaic dataset the improvement from no transfer learning to transfer learning from a general-purpose dataset was 9.78%. Furthermore, the ResNet50 architecture showed a high performance with better results than the UNet. The results indicated already that data balancing is an important topic. The study of invasive blueberry species, endangering sensitive wetland environments and black locusts invading into coastal forests, were two more application-oriented examples, with an easier problem definition: Classifying target species in a natural green environment. The ResNet50 architecture was used with transfer learning to detect black locust trees in an evergreen coniferous black pine forest with a 75% of True Positives (TP) and 9% False Positives (FP) while the detection of native trees achieved 95% TP and 10% FP. Detections of invasive blueberry bushes were performed with ResNet50, transfer learning and unfrozen weights with True Positive Values (TPV) of 93.83% and an Overall Accuracy (OA) of 98.83%. A refinement of the result masks reached a Dice of 0.624. Image analyses were performed to produce maps of blueberry location, distribution and spread in each study site, as well as bush height and area information. A preliminary study of different deep learning networks, transfer learning, the use of data augmentation and loss functions and settings were tested for the detection of invasive blueberry species. The challenge of the data was the imbalance, as invasive species had fewer individuals than natural occurring plants. In this study of state-of-the-art deep learning architectures the best results were obtained with the ResNeXt architecture (93.75 True Positive rate), and 98.11% accuracy for the Blueberry class with ResNet50; Densenet and wideResNet calculated similar results. The knowledge, gained with easy examples, was then applied to automatic tree species classification in natural mixed forest. This study provides an efficient and effective methodology to study forests and other natural environments, like wetlands, using different techniques: field surveys, image analyses and deep learning. Automatic generated results showed high accuracies and indicated the applicability of the methodology in different fields. Image analyses extract the most important information of aerial images, depending on the study focus. Field data captured a lot of information that could not be extracted from images, and therefore the methodology set provided new and important insights into forest environments.

(日本語)

地球上で 41 億 5,000 万ヘクタールの森林は、地表の広い範囲を占め、炭素循環の調整、生物多様性を保全、水供給など、経済的、環境的、社会的に不可欠なサービスを提供しています。人為的な攪乱や気候変動は、森林生態系における成長や回復力に影響を与えます。従来、森林に関する情報は、コストと労力のかかる毎木調査によって収集されてきました。毎木調査は通常小規模なフィールドプロットで行われ、そのデータを用いて森林レベルにスケールアップされてきました。しかし、日本に多く分布している天然生広葉樹林は非常に複雑であるため、小さなフィールドプロットでは調査対象の森林をあまり代表していないという問題が生じていました。

近年、UAV (Unmanned Aerial Vehicle : 無人航空機) が森林調査において用いられるようになってきました。UAV が搭載できる RGB カメラは集中的なキャリブレーションや前処理を必要としないため、データの収集と解釈が容易です。また、1m 以下の空間分解能で、林冠構造や葉の構造まで識別できるため、画像から樹種分類を行うことが可能です。

本研究では、現地調査、GIS を用いた画像解析、ディープラーニングという 3 つの方法論を用いて、日本の天然生混合林の樹種構成、分布、構造を特徴づけ、評価した。山形大学研究林 (YURF) の合計 13 サイトにおいて、フィールドデータを取得し、GIS アプリケーションを用いて画像解析を行いました。その結果、合計でフィールドデータからは 70 種の樹木が確認され、画像データを用いた深層学習からは 41 種のクラスが特定されました。林冠種を考慮した場合、画像解析とフィールド解析の結果は類似していましたが、下層植生と低木において両結果に差異が生じ、画像からはほとんど識別できませんでした。生物多様性の指標は、画像データとは対照的に、フィールドデータを使用した場合に高い多様度を示した。画像解析の結果では、シャノン多様性指数が 1.73~2.39 (均等性は 55%~82%) であったのに対し、2.14~2.76 (均等性は 71%~84%) であった。森林の階層構造に関して、画像からは下層においては識別できませんでした。溪畔林に設置したサイトは、実際にはオニグルミが優占する森林で、サワグルミが共優占種であったが、画像解析の結果によると、オニグルミ、カエデ属、ホオノキ、ミズナラが優占種であると認識され、優占度の順序が異なっていました。GIS ツールを使って画像を評価することで、データに隠れていた重要な側面が見えやすくなり、空間情報を簡単に抽出して解釈することができました。

一般的に、樹冠が密集していると個体数推定が困難になることが知られています。そこで、フィールドデータ、DEM、夏季および冬季の画像を用いて樹木個体数の正確な評価を半自動解析によりおこないました。その結果、冬の画像でカウントを行った場合に最も高い精度が得られましたが、小さい木や若い木のカウントは依然として困難でした。

次に深層学習技術を用いて植物種を自動的に分類するために、YURF の 13 のサイト、酒田市近郊の海岸林や、ドイツの湿地帯でもデータを収集しました。すべての画像は Metashape で処理され、オルソモザイクと DEM (Digital Elevation Models) を作成しました。まず、冬の画像で葉のある木とない木 (落葉樹と常緑樹) を分類するという簡単な例に深層学習を適用し、伝達学習と深層学習のアーキテクチャ (ResNet50 と UNet) の効果を評価しました。このアプローチでは、マルチラベルパッチ (MLP) 分類とセマティックセグメンテーションについても検討しました。その結果、MLP による落葉樹と常緑樹の分類で満足のいく結果を得るためには、転移学習が必要であることがわかりました。冬のオルソモザイクのデータセットでは、転移学習を行わなかった場合と、汎用データセットからの転移学習を行った場合の改善率は 9.78% でした。さらに、ResNet50 アーキテクチャは、UNet よりも優れた結果を示し、

高い性能を発揮しました。ResNet50 アーキテクチャを伝達学習とともに使用して、常緑針葉樹のクロマツ林におけるニセアカシアの木を検出したところ、真陽性(TP)が 75%、偽陽性(FP)が 9%であったのに対し、在来種の木を検出では TP が 95%、FP が 10%となりました。

本研究では、森林や湿地帯などの自然環境を調査するための効率的かつ効果的な方法論を、現地調査、画像解析、深層学習などのさまざまな手法を用いて提供することができました。自動生成された結果は高い精度を示し、さまざまな分野での方法論の適用可能性を示しました。画像解析では、調査対象に応じて航空画像の中から最も重要な情報を抽出します。フィールドデータは、画像からは抽出できない多くの情報を捉えているため、この方法論セットは、森林環境に関する新たな重要な洞察を提供するものだと考えられます。

以上より、本審査委員会は、「岩手大学大学院連合農学研究科博士学位論文審査基準」に則り審査した結果、本論文を博士（農学）の学位論文として十分価値のあるものと認めた。

学位論文の基礎となる学術論文

1. Kentsch S., Lopez C.M.L., Serrano D., Roure F., Diez Y., 2020. Computer Vision and Deep Learning techniques for the analysis of drone-acquired forest images, a Transfer Learning study. *Remote Sensing*. 2020, 12, 1287; doi:10.3390/rs12081287.
2. Kentsch S., Cabezas M., Tomhave L., Gross J., Burkhard B., Lopez C.M.L., Waki K., Diez Y. 2021. Analysis of UAV-Acquired Wetland Orthomosaics Using GIS, Computer Vision, Computational Topology and Deep Learning. *Sensors* doi.org/10.3390/s21020471.