

令和3年3月1日(月)

於・遠隔会議システムを使用したオンライン会議

## 令和2年度第二回国土地理院研究評価委員会

### 発言録

国土地理院

## 国土地理院長挨拶

院長の野田でございます。研究評価委員会の開催に当たりまして一言御挨拶を申し上げます。

委員の皆様方におかれましては、日頃より国土地理院が行っております測量行政に多大なる御尽力、また御指導、御鞭撻いただいております。この場をお借りして厚く御礼申し上げます。また、本日、年度末を迎えまして大変お忙しい中、この委員会に参加していただいております。誠にありがとうございます。

さて、前回の委員会で、来年度からの特別研究課題になりますが、SGDASの推計精度向上に関する研究についての御審議をいただいたわけでございます。先月（2月）13日深夜に、御承知のとおり、福島県沖を震源とする地震が発生したわけでございますが、このSGDASが発災後、16分後には推計結果を関係機関にお送りさせていただいております。今回の地震では、幸い大きな土砂災害は発生しなかったわけでございますけれども、深夜に発生した地震でございまして、こういった情報もたらされたということで、初動等の意思決定には大変重要だった、必要だったということを改めて認識したところでございます。

間もなく来年度を迎えるわけでございますけれども、来年度の予算では、もちろんこのSGDASの推計精度向上に関する研究を含めて、私どもが日頃行っております「国土を測る、描く、守る、伝える」という取組を引き続き切れ目なく行っていきたいと考えているところでございます。よろしく願いいたします。

さて、本日の議題でございますけれども、「AIを活用した地物自動抽出に関する研究」ということでございます。「国土を描く」という取組を効率化しようということございまして、平成30年度から5か年、すなわち令和4年度までの特別研究課題ということで取り組んでいるわけでございます。今年度、中間年ということございまして、中間段階の目標の達成度等を御審議いただくということにさせていただいております。皆様方におかれましては、様々な視点から忌憚のない御意見をいただきながら、御議論、御検討をお願いしたいと考えておりますので、よろしく願いいたします。

簡単ではございますが、以上で開会の挨拶とさせていただきます。本日はよろしく願いいたします。

## 平成30年度開始特別研究課題中間評価 1 課題【審議】

### AIを活用した地物自動抽出に関する研究

#### 1. 国土地理院：新規特別研究課題「AIを活用した地物自動抽出に関する研究」の説明

説明者：地理情報解析研究室 大野裕幸

##### 【スライド1-2】

今回はこちらの順番に沿って御説明いたします。約半分の時間を使いまして研究の位置づけと概要の部分を説明させていただき、その後、本研究の成果を順次説明していきたいと思っております。2.以降は、資料1-1にあります中間評価表の記載内容に沿った説明内容としております。

##### 【スライド3】

まず、研究の位置づけと概要について説明いたします。

##### 【スライド4】

本研究は、地図作成にAIを導入しようという研究であります。地図作成の工程というのは、昔からあまり変わっておりませんで、まず空中写真を撮影してステレオモデルを構築し、立体視をしながら写真の内容を判読して地物の位置を図化していきます。そして、図化した結果を地図記号に編集して、最終成果である地図情報が出来上がるというのが一般的な地図作成の流れになっています。

デジタル化が進んだ現在でも、地図作成というのは労働集約型の産業でありまして、多大な時間、すなわちコストを要するという課題があります。一方で、コンピュータの計算能力が高まってきて、AIの技術が様々な分野で実用化されるようになってきました。特に、画像認識の分野でのAIの発展が著しいという背景がありまして、地図作成にもAIの導入に関する検討が必要という観点から本研究が立ち上がったものです。それを受けまして、まずは判読・図化の工程を対象としまして、ブレークスルーを果たすようなAIを構築するということを目指して研究を進めてまいりました。今回、中間評価ということで、当初の想定とは大きく異なることになった過程の説明も含めまして、現在の取組状況を御説

明いたします。

#### 【スライド5】

実はAI、人工知能という用語に定義はありません。なので、本研究で使うAIとはどういうものかというのをまず御説明します。AIといいますと、言語認識のようなものを想像されるかもしれませんが、本研究が開発しようとするAIは、画像認識を行おうとするものです。具体的には、このスライドの左の画像のように、画像のピクセル一つ一つに何が映っているのかをカテゴリ分類していくというものでありまして、これは従来から土地被覆分類問題として取り組まれてきたものと同じと考えていただいて差し支えありません。その手段として、画像認識のためのAIを導入しようという点が異なります。そのAIを、入力画像と、それをどのように分類すべきかの正解を示す教師画像のセットを学習データとして与えて、深層学習という方法で判断の際の重みづけなどを学習させていくというものになります。

#### 【スライド6】

では、深層学習とは何でしょう。深層学習には、学習データ、目的関数、CNN (Convolutional Neural Network)、これは畳み込みニューラルネットワークの略ですけれども、この3つの構成要素が必要になります。学習データというのは、良質なものをある程度の数、用意する必要があります。何が良質に当たるかというのは議論のあるところでして、本研究の研究対象でもあります。2番目の目的関数というのは、いわばAIとしての判断基準に当たるものです。CNNのほうは、中身については時間の関係で説明は省略いたしますが、深層学習で使われるCNNというのは、一般に数百万から数億という大量のパラメータを有します。

深層学習の最初というのは、CNNのパラメータの初期値を乱数でセットすることが多いですけれども、入力画像をCNNに通して推論した結果と正解画像を目的関数というものに当てはめて比較をしますと、その違いが数値として算出されます。その値が0、すなわち推論画像と正解画像を同じにする方向にCNNのパラメータを補正していきます。補正したパラメータで次の画像を推論し、また正解と比較し、パラメータを補正するというのを延々と繰り返していきます。そうすると、次第に推論結果と正解の違いが小さく収束していけば学習は成功です。この過程が深層学習に当たりまして、最終的にCNNのパラメータ

の値という形で特徴量が自動的に学習されていくというのが深層学習の特徴です。本資料では、深層学習の結果として得られた学習済みのCNNのことをAIと呼ぶことにいたします。

#### 【スライド7】

そのような深層学習ですけれども、時間がかかります。本研究では、1週間単位で深層学習をすることも珍しくはありません。しかし、一旦学習済みのCNN、すなわちAIが出来上がりますと、すぐに事業など目的とする用途に使うことができます。使い方は簡単でして、AIに処理させたい画像を入力して推論させるだけです。

ところで、AIを使うことで従来と何が違うのかという疑問をお持ちかもしれません。AIを使うメリットというのは、処理速度と分類性能の2点だと考えています。分類に必要な着眼点や判断の重みづけというのは、大量の学習データを使って、CNNのパラメータとして学習済みですので、使うときはそれに従って画像処理するだけになりまして、処理速度は従来より非常に高速になります。また、深層学習による分類精度というのは従来よりも高くなることが多いようです。言い換えれば、地図作成にAIを導入することで、熟練の測量技術者がやっていることと同じことを極めて短時間で、かつ自動的に行えるようになるというのが従来との違いだと考えています。

#### 【スライド8】

こちらのスライドは、AIの分類性能をどうやって評価するかというのを示したものです。画像認識を行うAIの性能評価というのは、学習データとは別に用意した検証用のデータを推論にかけることで数値を使って客観的に行われます。その数値をスコアといいまして、推論結果と正解とを比較することで算出できます。たとえ異なるアプローチの研究であったとしても、同じ検証データを使ってスコアを算出することで、同じ土俵の上で性能の良し悪しを評価することができます。一方で、検証用のデータが研究中にころころ変わってしまいますと、スコアの相対的な比較ができなくなってしまいますので、本研究では最初に性能評価用の検証用のデータ作成を行いました。

スコアの算出方法には、F値——F1-Scoreともいいますが——とか、IoU (Intersection over Union) など幾つかあるのですけれども、本研究では、F値という2カテゴリ分類用の評価指標を使っております。

#### 【スライド9】

こちらはF値の算出方法の説明です。AIが抽出したところが正解である割合を示す適合率、それから抽出されるべき範囲の中で正しく抽出できた割合を示す再現率の2つの調和平均をとることでF値が算出され、0から1までの値をとります。F値が1ということは、推論結果と正解が完全に一致していることになります。

#### 【スライド10】

ここまでは位置づけと申しますか、AIとはどのような内容の説明をさせていただきましたが、ここからが本研究の概要の説明になります。本研究は、事前評価時と現在では研究のアプローチが大きく変わっています。こちらは本研究が着手時に想定していた研究の流れと研究要素の関係を示したものです。

#### 【スライド11】

スライド11とも併せて御覧いただきたいのですが、当初の研究計画では、地図と重なるように画像処理されたオルソ画像をベースにしまして、既存の地図情報や地形情報、その他のデータを地理座標上に集約したデータベースを最初につくりまして、そこから学習データや検証データを切り出すような形で自動的に作成することにしていました。それによって、大量に、かつ効率的にデータをつくれるだろうという目論見でありました。そうすると、AIの構成要素であるCNNと目的関数の開発に研究リソースを割けるなど。

#### 【スライド12】

しかし、思惑どおりにはいきませんでした。既存の地図情報と空中写真やオルソ画像の内容が、かなり広い範囲で一致しないということが分かったのです。その原因は、地図情報の更新時点と空中写真の撮影時期が一致していないためでありました。地図情報は新しい道路の開通など、どんどん更新されていくのですが、写真は撮影した時点のままということでした。

また、オルソ画像には本研究では許容できない量の位置ずれも見られました。特に、倒れ込みが致命的でした。オルソ画像というと、真上から撮影した状態になっているという

イメージですけれども、実際は、右側の写真のように、比高差の大きい建物のような地物では倒れ込みが見られまして、位置ずれが生じます。本研究は国土地理院の研究ということもありまして、位置精度を保つというのは、研究上、必須の条件となっています。右の写真の場合、倒れ込みによる位置ずれというのは、道路の車線2つ分くらいの量になっていますので、恐らく5 mを超えてしまっているでしょう。これでは本研究の成果にはなり得ません。また、倒れ込みによって隠れてしまった部分というのは、オルソ画像ではリカバーできませんので、お手上げになってしまいます。そのため、オルソ画像をベースとして最初に教師用のデータベースをつくって、必要なデータを大量に作成するという方法は断念せざるを得ませんでした。

#### 【スライド13】

その代替策として、オルソ画像ではなく生の空中写真を使うこと、そこから手作業で学習データと検証データを作成することとしまして、そのせいで研究全体におけるデータ作成の比重が相当に大きくなってしまいました。しかし、悪いことばかりではありません。この生の空中写真を使う方法に切り替えたことで、倒れ込みの影響の排除など幾つかの知見の発見につながりました。空中写真は同じ場所をいろいろな方向から撮影した画像であるという特性から得られたものですが、どういうものかというのは追って説明をいたします。

#### 【スライド14】

そのような”すったもんだ”の末に確定した研究の流れを整理しますと、このスライドのようになります。まず検証データを作成し、次にCNNと目的関数を決めてから、最後に学習データを作成して、AIとして組み上げることにしました。本研究は5年計画ですが、大ざっぱに①が1年目、②が2年目、③に着手したのが今年度からという状況であります。

#### 【スライド15】

一方で、本研究はAIの研究ということもありまして、達成目標を明確な数値目標として2つ設定しています。1つ目の数値目標は、公共測量の標準図式というのがありまして、そこに記載されている地図情報レベル2,500の項目のうち、40項目以上で人が判読した場

合に匹敵する抽出成功率の達成。しかし、これだとまだ曖昧さが残りますので、この抽出成功率の達成の部分でF値0.8以上としています。このF値0.8のレベル感というのは次のスライドで説明しますので、少しお待ちください。現在取り組んでいる40項目の内訳というのはスライド44に示しています。参考資料のほうです。

2つ目の数値目標は、1番目の目標をクリアした項目が占めるレコード数の合計が、抽出対象となる地物数全体の80%程度となることであります。

#### 【スライド16】

1つ目の数値目標で決めました各項目のF値が0.8という目標がどのくらいのレベル感かを説明したのがこちらのスライドです。研究着手当時、国際的なコンペといますか、コンテストが幾つか開催されていました。いずれもF値によるスコアを競うものでしたけれども、上のInria、これはフランス国立情報学自動制御研究所というところの略称ですが、ここが行っている建物抽出のコンペがあります。ここでは、当時の優勝スコアが0.7873でありました。

下側のDeepGlobe2018というのは、コンピュータビジョンの国際学会の一つでありますCVPR2018の一環でデジタルグローブ社が行った国際コンペでして、高分解能衛星の画像から道路、建物、土地被覆分類の3部門の抽出スコアを競うというものであります。こちらの優勝スコアは、道路はアリババのチームが0.65、建物は日本のパスコさんのチームが優勝されたのですけれども、0.74弱、土地被覆分類はやはりアリババのチームが0.53というのが当時の優勝スコアでありました。

これらのことから、F値0.8は、相当努力して達成可能かどうかという高いレベルの目標値として設定したものであることを主張しておきたいと思います。ちなみに、DeepGlobeは1年限りだったのですが、Inriaのほうは現在も続いておりまして、建物の最高スコアは現在0.8106となっています。

#### 【スライド17】

ここからが具体的な研究成果の説明となります。研究の流れに沿いまして、最初は検証データの作成から説明をいたします。



#### 【スライド18】

検証データというのは、AIの性能検証のためのスコア算出に用いるデータです。その分布というのは、AIの処理対象となる母集団の分布とできるだけ一致させることが望ましく、かつ算出されるスコアが評価値として有意となるだけのボリュームを有している必要があります。本研究では、学習データとは全く別に、独立した検証データセットを構築しました。本研究が目指すところは我が国の地図作成の自動化でありますので、母集団はすなわち日本全国となります。そこで、2007年度から2017年度に測量用のデジタルカメラで撮影された空中写真の地区は全部で549地区ありますが、その全てから機械的に1枚ずつ空中写真を選択しまして、検証用のデータセットを作成しました。この549という地区数は、先行研究よりも1桁多い数です。実は、このうち2地区だけ問題がありまして、最終的には547地区になったのですけれども、真ん中の日本地図のところに赤い点で表示したものが選出した画像の位置ですが、その地理的分布を見ますと、ほぼ全国万遍なく抽出できているかと思えます。

#### 【スライド19】

こちらが検証データの実例です。本研究では、40項目を目標として抽出に取り組んでいますが、そのうちの36項目がこの検証データの対象となっています。それぞれ正解値として項目別に人間が色分けして塗り分けたデータとなっています。諸元はこのスライドの右側に、塗り分け色は参考資料のスライド43に示しているとおりです。検証データには高い正確性が求められます。本研究では、元画像に空中写真を用いることにいたしましたので立体視をすることができます。がけや斜面、建物の一部など、立体視をすることで正確に範囲を同定することができ、作成に時間はかかりましたけれども、信頼性が高い検証データになっていると考えています。

#### 【スライド20】

次は、CNNと目的関数の検討についてです。2年目に行ったものです。

#### 【スライド21】

AIの構成要素というのは、学習データ、目的関数、CNNの3つでありました。検討に当たっては、まず道路の学習データを5,400枚作成しまして、学習データを固定しました。こ

の時点では学習データ作成の特別なノウハウは全く持ち合わせておりませんでしたので、単に無作為に画像を選択してデータを作成いたしました。学習データを固定することで、残る2つの要素のうち、片方を固定して、もう片方の内容をいろいろと変更することで、どのような条件がスコアにどう影響するのかを探ることができます。ほとんど順列組合せのようなことをやりまして、深層学習を繰り返して、スコアが上がる条件を探り出してきました。

#### 【スライド22】

その検討の結果、CNNには、U-Net構造と言われるのですが、そういうタイプのCNNを採用することにいたしました。当初はCNNではなくGAN (Generative Adversarial Networks)、敵対的生成ネットワークと言われるものを使っていたのですが、検討を続けるうちにF値が頭打ちになって、何をやっても上がらないという状態に陥った時期がありました。それで頭を抱えていたときに、深層学習をやっているとモニタリングのページが見られるのですが、それをボーッと眺めていたところ、GANの中で画像を発生させるGeneratorという部分があるのですが、その出来がいいことに気づきまして、このGenerator部分だけを取り出して深層学習させてみたところ、一気にF値が0.1くらい上がったみたいなこともありました。採用したのは、そのGenerator部分に使われていたCNNを基に、少しだけ修正を加えたものです。

目的関数のほうは、画素値のL1ノルムという、推論画像と正解の画素値の差の総和をできるだけ小さくするものを採用しました。多分画像認識の分野でL1ノルムを使っている人はそれほど多くはないと思います。この検討を通して、深層学習の条件などもよいものを見つけることができました。

#### 【スライド23】

次に、今年度から本格的な検討に入っております学習データの構成と、いよいよAIの構築について説明いたします。

#### 【スライド24】

学習データというのは項目ごとに作成しています。作成方法は至極単純でして、抽出性を意識して対象画像を選定し、こちらに示すように対象地物を項目ごとに決められた色

で塗り分けていくというものです。このようなものを一定の枚数をつくるごとに深層学習にかけてスコアを算出し、追加した学習データの効果を確認していきます。最初のうちは何千枚かつくったデータを全部捨てて、また一から作り直してみたいなこともありましたけれども、次第にコツとといいますか、F値を構成する適合率や、再現率の値の上がり方と学習データの構成にある程度の規則性のようなものがあることに気づいてからは、失敗は少なくなってきました。学習データを構成する画像の選び方には本研究のノウハウが込められています。

一方で、計算資源も人的資源も限りがありますので、できるだけ学習データ数は少ないほうがよく、無駄な学習データはできるだけ排除するようにします。

#### 【スライド25】

では、無駄な学習データとは何かということになると思うのですけれども、こちらは道路の学習データを無作為に作成した場合と、意図的に作成した場合のスコアの違いを示したものです。学習データ以外の深層学習の条件は全て同じにしてあります。この「意図的」の具体的な中身というのは本研究の重要な成果に当たるわけですが、同じ枚数でも、うまくつくればより高いスコアが得られるということが確認できました。F値が0.7を超えてきますと、なかなかF値は上がりにくくなってくるのですけれども、その中で10%違うというのは非常に大きな差です。

次に、F値0.77を記録した意図的5,400枚の学習データに対して、無作為に作成した5,400枚を加えて、1万800枚としたケースでは、さすがに0.8を超えるのではないかと思っていたのですが、実際はほとんど向上が見られないということも分かりました。

一方で、意図的に作成しますと、5,400枚に対し、僅か400枚追加した5,800枚の時点で、初めてF値0.8を達成することができました。つまり、F値0.7を超えるような領域では、もはやみくもに学習データの枚数を増やしても、ほとんどスコアの向上には貢献しないということでもあります。このことから、もしも学習データを自動的にたくさん作ろうとしていた当初の研究アプローチのままだったとしたら、大いなる時間の無駄を生じていた可能性があったということも分かったわけです。この「意図的」の部分のノウハウを得たことによりまして、良質な学習データとは何かという点について、ある程度つかむことができてきたと考えています。

ところで、先ほどはさらっと流してしまいましたが、「抽出性向を意識して選択」という

フレーズが気になって、“もやもや”されていたりしますでしょうか。本研究では、AIが2種類の抽出性向になるように学習データの作り方に差をつけています。

#### 【スライド26】

抽出性向の一つは、論理積を表すAND型のアイデアを実現しようとするAIです。別々の角度から撮影した空中写真の抽出結果の重なり合う部分だけを抜き出すことで、オルソ画像の場合に問題であった倒れ込みの影響をなくして、真上から見たときの断面形状、正射影とありますが、これを取得できるAIをつくらうというものです。そのためには、少々誤抽出があってもよいので、できるだけ漏れなく抽出する。すなわち、適合率よりも再現率の値が高くなるように誘導するという方向性で学習データを作成することになります。このタイプは、学習データの作り方としては実はあまり工夫せずに、たくさん学習データを作成すれば自然にこの傾向を呈してくるようです。

#### 【スライド27】

もう一つは、論理和を表すOR型のアイデアを実現しようとするAIです。これはAND型とは逆に、いろんな方向の空中写真から抽出した結果をマージすることで、抽出漏れを補完しようというアイデアです。これは生の空中写真を使うことで可能となるもので、オルソ画像ではこのようなことはできません。

こちらの例では、真ん中の画像では黄色の円内の建物がうまく抽出されていませんけれども、これに隣り合う左右の画像の抽出結果、特に右側の画像では円内の建物がうまく抽出されていますので、これをマージすれば欠落部分が補完されて、きれいに抽出できる可能性があります。ただし、誤抽出の部分もマージされますので、OR型のAIをうまく動作させるためには、抽出漏れは少々あってもよいから誤抽出の少ないAI、すなわち再現率よりも適合率優位に誘導する必要があります。こちらは意識して学習データをそろえないと、このタイプには育ちにくいようです。

#### 【スライド28】

こちらはOR型のAIのアイデアを実証した例です。長崎県の平戸市を撮影した連続する22枚の空中写真から道路を抽出した結果の一部を示しています。青い線の部分が、このう

ちの1枚だけの空中写真からの抽出結果でして、赤色の部分というのは、隣り合う空中写真の抽出結果からマージによって補完された部分を表しています。赤色の部分を含めることで全体として抽出漏れが少なくなって、この手法によって抽出率が大きく改善していることが見てとれます。これはまだアイデアの実証段階にあるものですが、今後これを熟成していく予定です。

**【スライド29】**

続きまして、目標の達成状況を自己評価いたします。

**【スライド30】**

本研究の成果は2つの数値目標によって評価されるものでありました。その数値を算出したものがこちらです。1つ目の数値目標をクリアした項目は現時点で5項目でして、目標達成率は5分の40で12.5%。各項目の学習枚数とF値の一覧は参考資料のスライド44に示しています。

2つ目の数値目標は、この5項目それぞれのF値とレコード数を掛け合わせた値から61.44%でした。各項目の占有率の一覧は参考資料のスライド45に示しています。

本研究では、当初は主要項目に絞って基礎的な検討を実施し、何度か作り直しも経験しながら学習データの選定方法を研究していました。それがようやく軌道に乗ってきて、作成項目を増やし始めたのが今年の10月のことですので、数値目標①の値は低いですが、想定範囲内と考えています。2つ目の目標数値は、主要項目に絞って検討を進めてきた結果が素直に反映された値かなと考えております。

**【スライド31】**

次に、現時点で想定している本研究の出口戦略について説明します。

**【スライド32】**

これまで説明してきた研究要素を組み合わせますと、空中写真を入力したら地図データが自動的に吐き出されるというAIが構築できます。特に、このシステムの場合の処理速度は、人が実施する場合に比べて圧倒的でありまして、数百人日かかるものが僅か半日で出来上がるというものになっています。しかも、現在構築しているこのシステムは、AIに処

理させたい空中写真を入力用のフォルダにコピーして、それから抽出したい地物の番号を指示するだけで、空中写真が何十枚であろうが何千枚であろうが、もう一気に処理してしまうというものになっています。ただ、まだ頭脳に当たる地物抽出性能はブレークスルーというほど高くはありませんが、今後この性能を高めていけば、どんどん抽出漏れなどのミスも少なくなっていくと考えています。そして、これが実用化できれば、我々はAIを導入する直接的なメリットを享受することができるようになります。また、このシステムを使って、空中写真からAIで抽出可能な地物は全て自動抽出できるようにするというのが本研究の最終到達目標になります。

#### 【スライド33】

ただし、こちらの図をよく見ていただくと分かりますが、まだその形が画像から抽出したままの形状なので、いわゆるジャギーで汚いのです。測量成果としての地図完全自動作成まで持っていくには、後続工程の地図編集部分をテーマとする研究開発が必要になってきます。

#### 【スライド34】

こちらは本研究の成果の現時点での活用事例です。位置精度の制約が緩く、スピードを求められる地物の抽出、すなわち災害対応を目的とするものについては特に相性がよく、実用化のレベルに到達していると考えておりまして、浸水のリアルタイム抽出システムに本研究のOR型のAIが既に使われております。

#### 【スライド35】

最後に、現時点の課題と今後の研究内容について説明します。

#### 【スライド36】

本研究の全体像というのはおおむね出来上がっていると考えていますが、1つ目の数値目標の達成状況が低いということは事実です。本研究の手法がマッチする地物では、F値は、こちらのグラフのように、学習枚数を増やせば徐々に向上していきます。今後、適切に学習データを増加させることができれば、各項目のF値の向上が期待できるということです。一方で、まだ多くの項目の学習データの数が1,000枚以下という状態ですので、今後

の学習データの作成が重要になると考えております。

#### 【スライド37】

一方で、被覆——これはコンクリートなどで覆われた斜面のことですが——とか、人工斜面（土がけ）のような、画像だけからでは抽出が難しく、高さも考慮しないとうまく抽出できそうにない項目があることも分かってきていまして、全く新しい抽出手法を開発する必要が生じています。特に土がけというのは、地図データベース全体の16%を占めていまして、これの抽出方法を発見しないと2つ目の数値目標の達成は困難であります。このような、いわゆる立体視をしないと判読できない地物の抽出では、AND型のAIがキーになる可能性があると考えています。

#### 【スライド38】

また、研究内容を少し修正しています。当初計画では先進光学衛星の画像を使うことにしていましたが、残念ながら打ち上げが来年に延期となったということで、研究対象からは外すこととしました。また、レーザスキャナによる3次元点群データは、地図情報レベル500から1,000の地物に対応するものが主であることが分かり、オーバースペックということで、こちらも研究対象から外しています。

これまでの研究で研究の大枠が固まりましたので、今後は各項目のトレーニングデータづくりに集中し、引き続きブレークスルーの実現に向けて努力していきたいと考えています。

#### 【スライド39】

参考文献はこちらです。

私からの説明は以上です。

## 地理分科会概要説明

○委員長 どうもありがとうございました。

それでは、委員会に先立ちまして行われた地理分科会での議論について、分科会主査のほうから御説明をお願いいたします。

○地理分科会主査 研究開発の必要性に対する地理分科会としての総評と幾つか出されま

した意見について、まとめて御報告させていただきます。

まず、地理分科会は2月10日の14時から15時34分まで行いました。当日は、地理分科会委員3名全員出席し、意見を交わしました。次のとおり御報告いたします。

まず、研究そのものとしては、この研究の最初で説明がありましたように、オルソではなく空中写真そのものを使うという、当初の想定とは異なっていたわけですが、逆にそれがその後の展開には有利に働いて、内容的には順調に行われていると判断しています。

また、現状では、後半のほうで説明がありましたように、当初の目標スコアへの到達が途上ではありますが、まだこれは中間段階であって、到達するための方法や問題点が明らかになっていることから、これは達成できるのではないかと判断しました。

また、ただいまの発表ではあまり触れていませんでしたが、事業化やこの研究成果の実装化というのは、ここで立てられている研究目標以上のスコア、適用性が求められるわけですが、それに関しても、到達するための方法について示唆を得ているということでありました。

個々の内容について、特記的な議論や意見について少しお話しさせていただきます。先ほどの到達目標のF値0.8というスコアですが、このチャレンジングな目標に対して、今後、効率的に教師データをつくり学習させるということで、到達目標の達成を目指すという明確な方向性が見えており、今後もしっかり取り組んでいただきたいという意見です。

また、最後のところで述べられた浸水状況の把握といったようなことに非常に親和性があるということでした。成果公表のところで、浸水状況のリアルタイムの把握に関して、この研究が非常に応用性が高いということも分かって、地図作成分野だけではなく、教師データを置き換えることで様々なことに応用ができるということを示しているのではないかと判断しております。

また、成果の公表は、まだまだ進んでいない部分がありますけれども、現在、AIの研究成果は、多様な研究者がデータセットを用いてCNNをブラッシュアップしていくという事例がこの分野で見られるということですが、国土地理院で上げた研究成果も、地理院作成のデータセットとして公表されていくことによって、国土地理院で行う地図作成だけでなく、地図作成分野全般に大きく貢献できる、社会貢献ができる研究であるということが期待されるだろうという意見です。

そして、最初のところで述べましたように、実装にはまだまだ時間がかかって、この研



究が終わった後に、新たに実装に向けての研究が始まるのだと思いますけれども、その実装へ向けての方向性や今後の公表の方向性なども議論いたしまして、この研究を継続してしっかり進めてほしいという意見でありました。

ということで、分科会で議論された内容の概要と意見について述べさせていただきました。以上です。

○委員長 どうもありがとうございました。

それでは、ここから今日御参加の委員の皆様から御意見や御質問をお願いしたいと思いますが、分科会での主査以外のお二人の先生から何か追加のコメントはございますでしょうか。よろしいでしょうか。

○委員（地理分科会構成員） 1、2 特にありません。

## 各委員との質疑応答

○委員長 それでは、分科会以外の委員の先生方からいろいろ御意見をいただきたいと思っています。最初に、このテーマが3年前に提案されたときに、5年間というすごく長いスパンになるので、中間発表はぜひやっていただきたいということを当時申し上げていたと思いますし、もう一つは、AIの流れだとか、いろんなものが恐らく研究をする上で変わってくるだろうから、そこは無理だと思ったら、いろんな方向性をドラスティックに変えていくこともあっていいのではないかということ、私はたしかコメントした記憶がございます。そういうことも含めて、委員の先生から今発表があった内容について、御意見、コメントをいただければと思います。対面では、ここでどなたかが手を挙げられればパッと当てられるのですが、今日は前回と同様ウェブでございますので、順次お話をいただくということで御了承をいただきたいと思っています。

それでは、順番にコメントをいただきたいと思っています。どなたからでも構いません。よろしく申し上げます。御発言のときにビデオをオンにさせていただくと、どなたかが分かりやすいので助かります。

## 委員1からの質問

○委員1 御説明いただきまして、ありがとうございます。ちょっと質問ですがけれども、最後の御説明の中で、高さが関与する地物の抽出は苦手だという話があったのですがけれども、地図に落とす場合には大抵3Dではなく2Dなので、立体視が苦手ということが具体的

にどのような影響を及ぼすのか、私は理解できていなかったの、教えていただければと思います。

○説明者（国土地理院） 現在、我々が作成しているAIというのは、厳密に言えばカメラパラメータもとるのですけれども、基本的には入力として空中写真のみ、すなわち今画面に映っているような色分けしていないほうの画像のみを入力して、そこからいろんな地物を見分けていくというAIをつくっているわけですが、特に問題視している被覆と土がけに関しましては、写真のみからではどちらが高いかというのが全然分からないのです。例えば、右側で道路が映っていると思うのですが、その左側に本来であれば土がけとして、緑色に塗っている部分が土がけですよということで教師データでは示しているわけですが、これの画像だけを見ると、高さが分からないので、荒地だと判断してしまうのです。立体視の場合は、右目と左目の2枚の画像を組み合わせることで視差が生じて、そこに比高差があるということが分かるわけですが、今のところ入力は1枚の画像のみとしていますから、それが現状では把握できないわけです。なので、できれば1枚の画像からそれを分かるようにする手法を開発したい、する必要があると捉えているということになるわけです。

では、右目と左目の2枚の画像を一緒に食わせればいいのではないかとお考えになるかもしれません。当然そのような発想に向かうことは承知しているのですが、2枚の画像を入力して立体視をさせようとした場合、実は条件が非常に厳密になってくるのです。結局、位置を合わせないと視差の情報が取れないので、2枚の画像の位置を厳密に合わせないと、推論させるときに全然そういった情報が得られないということになってしまいます。トレーニングのときにそれでやってしまうと、使うときにも同じような条件で情報を食わせないといけませんので、実際に使う場合に非常に敷居の高いものになってしまうということで、できるだけ使う場合に使いやすい1枚の画像のみでこういったことが抽出できることを目指して手法を開発したいということになるわけですが、現時点では、まだこれをどういうふうに解決すればよいかはノーアイデアというのが正直なところですが、これから検討していかないといけない状況になっているというところであります。

○委員1 よく分かりました。ありがとうございます。

## 委員2からの質問

○委員 2 F値というのでずっと議論されていて、それはふだんあまり使わないので、感覚的によくつかめなかったのですけれども、スライドの9にその定義が書いてあったのですが、要するに、PrecisionとRecallが両方とも9割ぐらいのときにF値が8割ぐらいになりそうな気がいたします。ということは、10個のうち1個ぐらい間違えるのが最高性能ということかなと思って、そんなものでいいのかなと素朴に思ったので質問しました。

○説明者（国土地理院） まず、F値というのは、おっしゃるとおり、適合率と再現率、この絵で言うと、四角の枠内の青い円内のところのピンク色の部分の割合と、赤い円内のところのピンク色の部分の割合とが重要になってくるわけですが、その2つというのは通常トレードオフなのです。適合率を上げれば再現率が下がるし、再現率が上がれば適合率が下がるという関係にありますから、どちらか一方だとなかなか評価が難しいということで、調和平均という形をとって、F値という形で性能を数値で表しましょうという種類のものです。

実は、0.9、0.9だとF値は0.9になるのですけれども、例えば両方とも0.8ですと、調和平均なので0.8になりますし、片方が0.9で、片方が0.72ぐらいでも0.8ぐらいになるのです。0.8という数値に関しては、そもそもこの手法自体で世界中の皆さんが取り組んでいらっしゃるわけですが、いかにして取り組んでも、なかなか0.8という数値はクリアできない数値であるというのが1つあるのですがもう一つ、説明では申し上げませんでしたけれども、我々はU-Net構造というCNNを使っていると説明させていただいたのですが、このU-Net構造というのは、AIとしての絶対的な性能を見たときに、必ずしも高いものではないのです。同じ学習データを使っても、もっと高いF値が出るCNNというのは実はほかにもありまして、それを使うと、もっと高いものを数値として得ることはできていると思っております。

ただ、我々の研究成果の主なものは学習データセットなのです。地図作成用の学習データセットを我々が世の中に提供して、それを同じ分野の研究者の皆さんに使っていただくと、少なくともF値が0.8というのは出すことができますというデータセットとして作成しているものになりますので、何が言いたいかというと、とにかく高いF値をただ求めるということではなくて、確実に0.8——0.8というのも非常に高い目標値ではありますけれども、どういうCNNを使ったとしても0.8を出すことができるデータセットという形で、世の中に成果を研究成果として提供しようということを目指してやっています。先ほど地理

分科会主査より、地理分科会のときの事業に関する部分には触れなかったけれどもということでおっしゃいましたけれども、事業に使うとなると、確かに0.8ではまだ低いと思います。ただ、0.8から事業に使う新たなステップに進むためのベースとしてのスタートラインとしては、少なくともF値0.8が出るようなデータセットが必要ということで、我々は、それをどうすればより少ない学習データの数で達成することができるのかということを目指して研究を進めてきているところですので、そういう位置づけでF値0.8という設定をしているということをお願いさせていただきたいと思います。

○委員2 どうもありがとうございます。

### 委員3からの質問

○委員3 1つ前の質問に戻ってしまって申し訳ありません。スライド37、地形差があるときのものにまた戻ってしまうのですけれども、ここで先ほど1枚の写真からということですが、もともと高さの情報は1枚の写真には含まれていないと思うのです。立体視しなければ出てこないものだと思うのですが、幾らAIでも、もともと入っていないものから高さの情報は出せないのではないかと危惧するのと、もう一つ、ここに新たな抽出する手法の開発が必要と書いてあるのですけれども、今この研究の中間評価ですが、研究の後半でこれに取り組むということなのか、それとも、また別建ての研究テーマを設けてそちらでやるのか、どちらなのでしょう。

○説明者（国土地理院） 後のほうの質問から先に回答いたしますと、この研究の中で挑む予定であります。目標のクリアに挑むためには、それを解決する必要があるという観点から、あと2年間でこれを解決しようと考えているところであります。

前のほうの御質問ですけれども、おっしゃるとおり、1枚の写真からでは高さの情報というのは読み取ることはできません。ただ、AND型という形で我々が解決している——まだこれは実証ができておらず、アイデアの段階ですけれども、1枚の画像から抽出した結果を正射変換すると地図上に抽出範囲を落とすことができます。隣り合う別の空中写真から同じように抽出した範囲を地図上に落としますと、これは東京スカイツリーの場合ですけれども、このように重なるわけです。ここまで重ねることができれば、視差の情報が取れるかもしれない。例えば、これだとタワーの一番てっぺんの部分が、上の突起のような部分と下側の突起のような部分の2つに分かれてしまっていて、その差の分だけ要は比

高差があるということになるわけですが、これをがけの場合にうまく出ないだろうかと。右の写真で撮ったがけの部分と、左から見たときのがけの部分をもっと組み合わせ、それが同じがけを撮っているということを認識させることができれば視差が取れるだろうと今、漠と考えているところですが、AND型のAI自体の実証がまだできていないということもありますので、その手法で果たしてうまくいくかどうかは、まだはっきりとは分からないのですが、いずれにしても、1枚の画像からでは難しいので、AIで抽出した結果を後処理という形で組み合わせることによって、高さなど、そのほかの情報も抽出できるのではないかというアイデアを持っておりますので、その形で検討させていただきたいと考えているところであります。

○委員3 分かりました。どうもありがとうございます。

#### 委員4からの質問

○委員4 非常に興味深く拝見いたしました。お伺いしたいことがございまして、先ほどF値0.8というお話をされていて、それからスライド32に「ただし、測量成果とするには、さらに形状を整形する等の処理を加える必要があり、現時点で『測量成果としての地図』の自動作成まで到達しているわけではない」というコメントもやっぱりなされていて、全てAIで測量成果としての地図の自動作成を目指されるのか、それとも、AIではある程度のところまでの到達点を見込んで、それ以降はさらに加工を加えていくとか、そういった対応も考えられるのでしょうか。どちらでしょう。

○説明者（国土地理院） 私は研究している立場なので、研究者としては、やはりそういう後処理の部分まで含めてAIで処理できるようにしたいと思って取り組んでいます。ただ、我々は、今、研究室レベルでやっているものと、それから事業レベルがまだその次の段階としてありまして、研究室レベルでは0.8を目標でやっているわけですが、事業レベルから見ると、0.8だと、とてもじゃないですけども、話にならないレベルなのです。そこからさらに何らかの工夫を加えて、私は過去の経験上、0.97というラインが事業としてコスト的に成り立つレベルというのですか、AIで実施した結果を人が補完するというこの人にかかるコストと、AIでいろいろな後処理をしてかかるコストが、0.97を境に逆転する。

○委員4 そうなのですか。

○説明者（国土地理院） 昔、地形図が画像だった頃に、それをベクトルデータに変換するということをやっている、全国を取り組んだときに、残り3%というのが事業として発注にのっかるときの判断基準だったのですよ。だから、0.97を達成するところまで行けば、AIのほうに切替わりが始まってくるだろうと思っています。研究としては確かに、この手法だけをとってみれば0.8をクリアするのは非常に大変ですけども、ただ、そこに、また先ほどのOR型とかAND型とか、抽出した結果を使ってさらに何らかの組合せをして、実は道路と建物などを演算しながらということまで考えているのです。そうすると、個々の項目だと0.8しか出ないのだけれども、ほかのところと組み合わせると、さらに実際のF値は上がっていくのではないかと考えていまして、例えば道路で見ると、Fだと調和平均なので0.8ですが、Precisionが例えば0.88で、Recallが0.72とかというケースがうちの場合は多いのです。そうすると、誤抽出が少ない。その点をうまく捉えると演算ができるんですよ。例えば、建物の誤抽出は道路上には出ることがあるけれども、道路の誤抽出が建物上に出ることはないという条件が分かっていると、建物から道路の抽出結果を引いてあげれば、計算上さらにPrecisionが上げられるわけです。そういったことをこれから残り2年の間に、個々の地物、まだ僅か3種類、5種類ぐらいしか0.8をクリアしていないですから、あまり偉そうなことは言えないのですが、残りの地物に対してもそれをクリアできて、かつ、それがPrecision優位なのか、Recall優位なのかということを見極めた上で、いろいろ組合せをしていけば、どれだけ0.97という目標値に到達できるのかにチャレンジをしたいと。

ただ、研究計画上は、その部分はこの5年間の次の研究の課題と捉えていましたので、後続の地図編集の過程を経る必要があるということをごういったところに記載しているわけですけども、研究者としては、できるだけ前倒しで到達したいと考えているところがあります。

○委員4 F値0.8はすごいと思います。

○説明者（国土地理院） かなり難しいと思います。

○委員4 対象にしているものはこういうものではないのですが、別のところの画像処理のAIで、0.8なんか出ていない論文が普通にいっぱいあるので、すごい値だなと思いつつ伺っていて、0.97といたらさらにもものすごいなと。

○説明者（国土地理院） 0.97は、さすがに現状では難しいですけども、0.8のレベル感がなかなか地理院内だと理解していただけないものですから、ありがとうございます。

○委員4 むしろ地理院の方に聞いていただきたいのは、F値0.8に達していない論文は、人工知能のこういう画像なんかを対象にしていっぱいありますよ。情報系とか人工知能なんかの学会では普通にあります。それはコメントでしたが、さっきからF値0.8とおっしゃっていたので、それが私はすごく気になっていたのです。

○説明者（国土地理院） 0.8は、苦勞はしましたけれども、1回クリアしてしまうと結構当たり前というか、頑張れば出るぞという感じにはなってきます。

○委員4 いや、当たり前ではないと思いますよ。0.97といたら「えー」とか本当に思います。

○説明者（国土地理院） 0.97もそんなに夢物語ではないなという感じにはなっておりません。

○委員4 それは研究成果をこれからお伺いするのを楽しみにしています。ぜひ頑張ってくださいたいのと、もう災害対応で結構実用化されているじゃないですか。私は、そこもすごいと思います。

○説明者（国土地理院） まだ実用化までは行っていなくて、試験運用の段階ですけども、あと1年ぐらい検討した上で実用化に持ち込む例ではあります。

○委員4 災害のときに、直後で一番そういうデータがみんな欲しいのですけれども、なかなかなくて苦勞しているという話は聞くので、ぜひ実用化を特に災害対応は早めにやっていただけると、とてもありがたいと思います。最後はコメントでした。

○説明者（国土地理院） ありがとうございます。

## 委員5からの質問

○委員5 大変興味深く拝見しました。質問です。本研究の位置づけと伺いますか、日本国内でほかのこのような地図でのAI作成をやっている研究と、世界中のほかのところでや

っている地図自動作成との比較で、本研究のステージというか、レベルというのはどれぐらいなのでしょう。それから、ほかの企業なり研究所が行っている研究を取り入れると  
いうか、比較もしているのでしょうか。

○説明者（国土地理院） ちょっと音声聞き取りづらかったので、こういう趣旨の御質問かなということも斟酌（しんしゃく）してお答えしますが、AIの研究というのは、多くの場合、ベースとなるデータセットが公開されていて、それをみんなが共通して使うことでCNNとか、目的関数とか、要は頭脳に当たる部分の改良を行って、こういうアイデアのCNNをつくりましたということで公表して、みんながさらにそれを導入してという形でブラッシュアップしていくというのが大体研究のモデルになっているのですが、その中で重要になってくるのは、基になるデータセットになります。実は地図作成の分野で、メジャーなというか、みんなが共通して使っているようなデータセットというのは、私はまだないと思っているのですが、事例として自動運転の分野で、車で人が運転席から見ているビューがございます。その中から、どこに人がいてとか、どこに車があってということと同じような手法で抽出するという分野があるのですが、例えばそれに関しては、ドイツのメルセデスが、CityScapeという非常に信頼性が高く、かつ画像の枚数も多い、データとしてしっかりしているというものをドーンと世界に公表されたのです。

そのような基となるデータセットで、しっかりしたものが公表された分野というのは、AIの研究が物凄く速く進みます。私は最近、自動運転の技術がどんどん進んでいるというのは、そういうことも要因の一つとしてあると思っていますので、これは研究者としての希望ですが、そういったモデルをこの地図作成の分野にも実現したいと考えています。我々はどちらかというとCNNの専門家ではなくて、やっぱり地図データの専門家ありますので、地図作成の分野でベースとなる信頼性の高いデータセットをAI研究の分野に公表することで、地図作成もしくは地理空間情報を画像から自動抽出するという研究分野をより発展させていきたいと考えているところであります。もし御質問の趣旨が違っていたら申し訳ありません。

○委員5 音声聞き取りづらくてすみません。今の御回答はよく分かりました。私が聞きたかったのは、本研究が他研究との比較でどれぐらいの位置にあるのかということです。簡単にいえば、誰も研究していないことを独自でやっているのか、ほかのところが研究しているとすれば、もっとすごい上に行くところがあるのか、そういう相対的な位置づけを



お聞きしたかったのです。

○説明者（国土地理院） 分かりました。ちょっと趣旨が違っていたみたいで、すみません。

この分野というのは、道路とか建物といった主要な地物は結構多くの方が取り組んでいらっしゃる。ただ、我々が取り組んでいる地図作成で、40項目というのが目標になっているのですけれども、特に主要な、我々が1年目、2年目で限定して取り組んだような項目以外の分野というのは、皆さんが取り組んでいらっしゃるという論文なんかあまり見たことはないです。1つは、皆さんが取り組んでいらっしゃる主要地物について、今、我々のF値がどのぐらいのレベル感にあるのかというのを捉えながら、そのノウハウなりをほかの人が取り組んでいない地物、だから、あまり数が多くない地物というのですか、そういったものにどのように生かしていくのかを研究するのが、日本の地図データをたくさん保有しているといいますが、写真をこれだけ大量に持っているのは日本では地理院ぐらいだと思いますので、そういう機関の役割として果たしていきたいと考えているところであります。

○委員5 分かりました。ありがとうございます。

#### 委員6からのコメント

○委員6 コメントなのですが、研究成果をお伺いしていると、口頭発表を人工知能学会の全国大会でされていますよね。大野さんの御研究を人工知能学会とか情報処理学会とか、そちらに投稿されたり、もうちょっと発表されると、いかにさっきのF値0.8がすごいかというのが証明されるといいますが、そういうことにもつながると、人工知能の分野でも評価されるべき研究だと思うので、そういったところにも研究成果を発表していただけるといいと思います。最後はコメントでした。

○説明者（国土地理院） ありがとうございます。努力いたします。

○委員6 多分そちらはいい評価をされますよ。ありがとうございます。

#### 委員7からの質問

○委員7 地理分科会の際に気がつかなかったことではございますが、最初のほうで、

国際コンペで高解像度衛星によるものとの比較をされていたのですけれども、衛星からだと倒れ込みというのは少なくなるのかなと思うのです。そうすると衛星のデータをこのシステムにプラスして精度が上がるとか、そういうことはできるのでしょうか。素人なので申し訳ありません。

○説明者（国土地理院） 衛星画像の場合に倒れ込みがないかということ、実は必ずしもそうではないのです。カメラの場所が空中写真とは違うので、直下がきれいに撮れば、かなり倒れ込みの少ない画像が撮れるとは思いますが、衛星画像の問題点というのは、実は地上画素寸法です。我々が取り組んでいる空中写真の地上画素寸法は20cmというのが基本になっているのですけれども、こちらで表示しているDeepGlobe2018で使われたのは30cm級の衛星画像、衛星の名前がすぐ出てこないのですが、衛星画像で地上画素寸法が一番小さいのは30cm級にとどまってしまう。一方、2,500のレベルの地図をつくる際に必要な画像の地上画素寸法は20cmと規定されていることもありまして、30cmの状態のもので2,500レベルは厳しいなということです。当初は、本当であれば今年打ち上げ予定であった先進光学衛星、これは80cm級の衛星ですけれども、この衛星画像を使うということも想定はしていたのですが、その場合は2,500ではなくて、2万5,000の地区ということで考えておりましたので、この場合、80cmというのがジャストサイズといいますか、逆に今度は30cmの衛星画像を持ってくると非常にコストが高くなってしまったという問題が生じますので、80cmはバランスがよかったですけれども、残念ながら衛星の打ち上げが1年延びてしまったということなので、衛星画像と組み合わせてという方式は取れなくなってしまったところであります。

○委員7 ありがとうございます。20cmと30cmだとやっぱり違うということですね。

○説明者（国土地理院） 規定上の問題になるのですけれども、AIを使うと、人が同じ写真を使ったときよりも多分位置精度はよくなるのではないかと考えているのです。ただ、規程上、どうしても2,500レベルの測量成果をつくるためには、その元画像としては1万分の1レベルの空中写真、すなわち地上画素寸法20cmの画像が必要ですよということになっています。最終的な本研究のアウトプットは測量成果に匹敵するものをつくるのが目的ですので、その点から30cmだと入力画像としては足りないかなというところでもあります。

○委員7 ありがとうございます。

## 講 評

○委員長 ありがとうございます。ほかの委員の先生方はございますでしょうか。よろしいでしょうか。

それでは、これから委員だけで講評についての議論をしたいと思います。今、3時25分で、一応20分という予定になっておりますので、ちょっと予定の時間を超えますけれども、3時45分ぐらいまでには終わらせたいと思いますから、委員の皆様はそのままお待ちください。国土地理院の皆さんのほうは、スピーカーの音量をミュートにして席を離れていただいて、45分になりましたらお戻りください。委員のほうの議論が早く終わった場合は、その間、休憩ということにさせていただきます。

それでは、委員の方はそのままお待ちいただいて、地理院の方は離席をいただいて、3時45分にお戻りください。

(委員のみによる審議)