

Neural Network Console クラウド版

ネットワーク解説

-領域抽出編-

ソニーネットワークコミュニケーションズ株式会社

概要

本ドキュメントではNeural Network Console(NNC)にある領域抽出のサンプルプロジェクト([tutorial.semantic segmentation.binary semantic segmentation](#))のネットワークを解説します。

サンプルプロジェクトをベースに自らネットワークを変更して精度改善を図るために、現状のサンプルプロジェクトのネットワーク構造を理解したい方を読者と想定しております。

これから領域抽出を始めようという方は、領域抽出の一連の流れを解説した[スターターガイド-領域抽出編](#)をまずはご確認ください。

本ドキュメントでは各レイヤー※¹がネットワーク全体でどのような役割を担っているかに焦点を当て説明していますので、ネットワークの中で用いられているレイヤーの具体的な機能は[レイヤーリファレンス](#)をご確認ください。

※¹ レイヤーとはDeep Learningでネットワークを作成するための関数で、NNCに限らず一般的なもののため、それぞれの詳細な仕組みなどは入門書などでも確認することができます。

目次

領域抽出のサンプルプロジェクトにはネットワークとデータセットが含まれています。

ネットワークの全体像を理解いただくために、まずはデータセットとネットワークの入出力を解説します。
そのあとにネットワークの詳細な構造を解説していきます。

1

データセットとネットワークの入出力

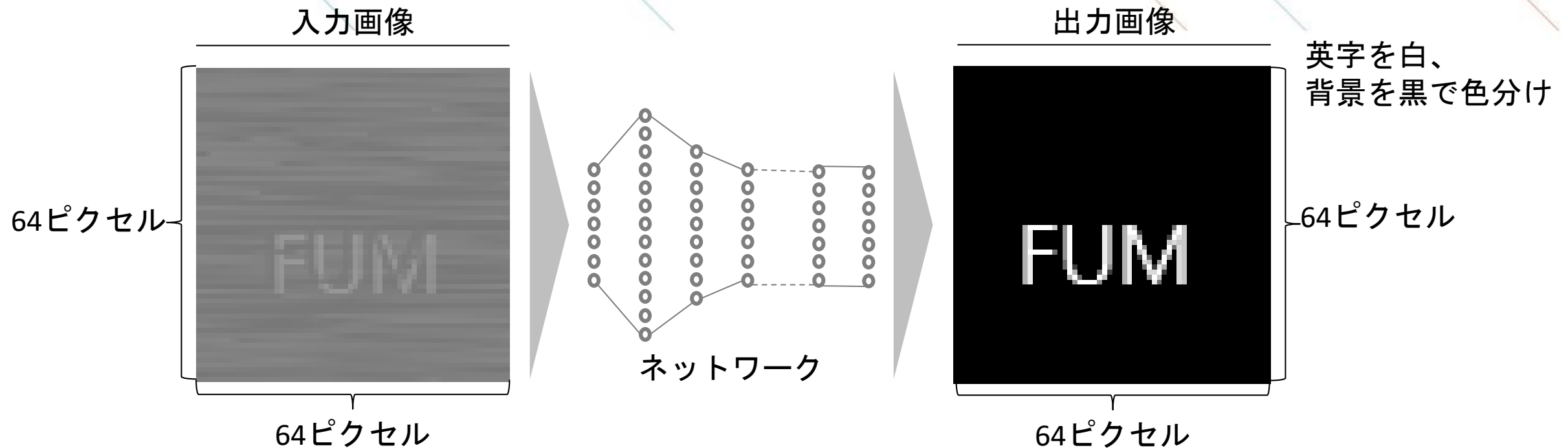
2

ネットワークの構造

データセットとネットワークの入出力

領域抽出のネットワークでは画像を入力とし、該当する部分をピクセル単位で分類した画像を出力します。今回のサンプルデータ(synthetic_image.binary_segmentation)では、モノクロ画像を入力データとし、英字部分を白、背景を黒で色分けした分類画像を出力します。

具体的にはヘアライン加工された金属板の上に印字された英字部分を認識するデータセットです。



目次

ここまでネットワークの全体像を理解いただくために、データセットとネットワークの入出力を解説をしてきました。ここからはネットワークの構造を解説していきます。

1

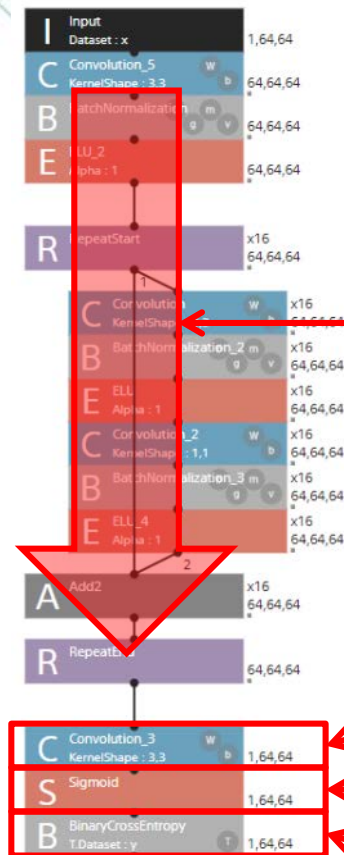
データセットとネットワークの入出力

2

ネットワークの構造

ネットワークの全体構造

サンプルネットワークはConvolutional Neural Network(CNN)構造を繰り返すことで、入力の画像サイズを変えずに特徴量を抽出しています。ネットワークの最後の部分で、複数枚の特徴量マップを1枚に集約し、ピクセル値を0~1に変換することで、文字部分かどうかの確率値を出力しています。

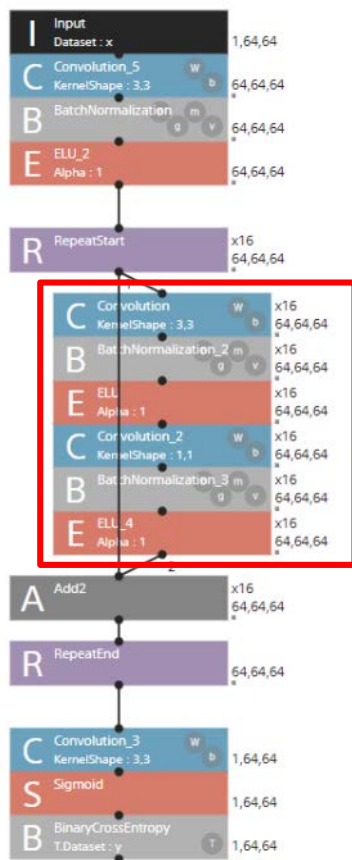


64x64の画像サイズを変えずに、CNN構造を利用し、特徴量を抽出
64枚の特徴量マップを生成し、ネットワークの中でアップデートを実施

64枚の特徴量マップを1枚に集約
各ピクセル値を0~1に変換
正解の出力画像(二値化画像)との誤差を計算

Convolutional Neural Network構造

サンプルネットワークはCNN構造をベースにしています。Convolutionレイヤーの詳細設定のBorderModeをsameに設定し、画像サイズが変化しないようにしています。OutMapsの枚数を増やすことでモデルが複雑になり、高精度が期待できます。一方で計算量が多くなるため注意が必要です。その他のCNN部分の改良については、[CNNの動画解説](#)をご参照ください。



基礎となるCNN構造

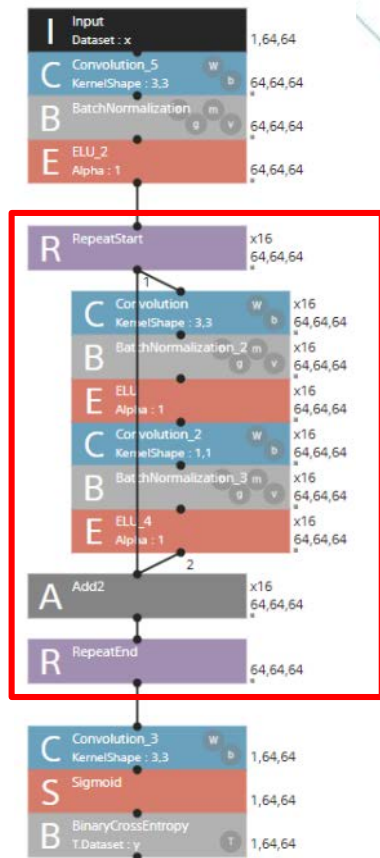
Convolution	
Name	Convolution
Input	64,64,64
OutMaps	64
KernelShape	3,3
BorderMode	same
Padding	1,1
Strides	1,1
Dilation	1,1

← 特徴量マップの枚数を指定

← 画像サイズが変わらないよう設定

Repeat機能を用いた繰り返し

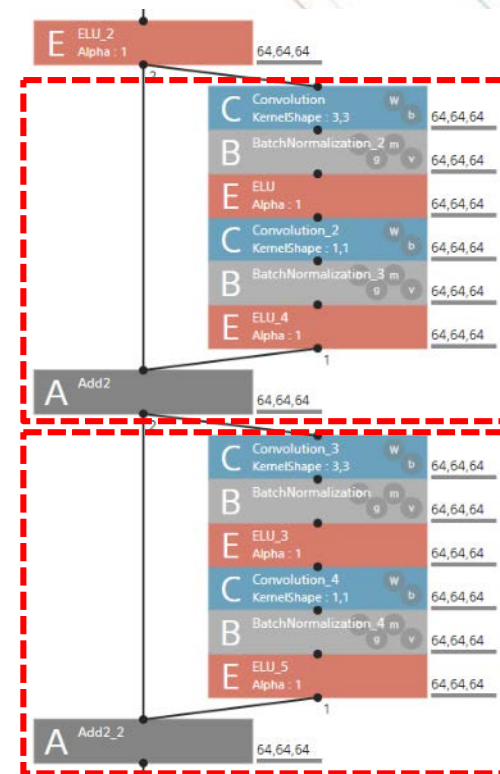
サンプルネットワークはRepeat機能(RepeatStart、RepeatEnd)を用いて、CNN構造を複数回繰り返しています。RepeatStartのオプションで繰り返し回数を設定しており、繰り返し回数を増やすことでモデルが複雑になり、高精度が期待できます。一方で計算量が多くなるため注意が必要です。



Repeatによる
繰り返し

R RepeatStart	
Name	RepeatStart
Input	64,64,64
Times	16
ID	RepeatStart
Output	64,64,64

展開すると



RepeatStartから
RepeatEndの中の部
分が連続で繰り返し

Timesで繰り返
し回数を指定



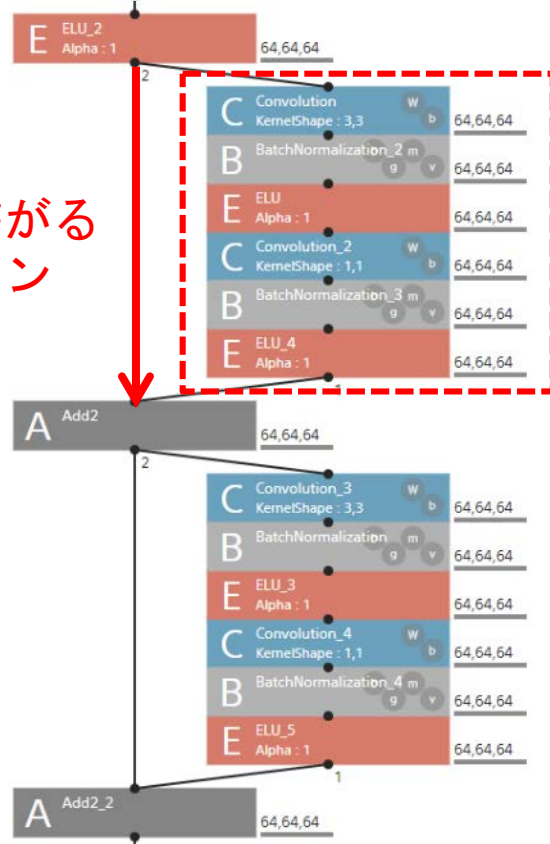
Residual Network構造

サンプルネットワークでは、単にCNN構造を繰り返すだけでなく、ショートカットコネクションを取り入れたResidual Networkを取り入れています。

Residual Networkを用いることで、CNN構造を多層化した際に学習が難しくなる問題(勾配消失問題)を解決することができ、さらに高い予測精度を実現できるということが知られています(次頁参照)。

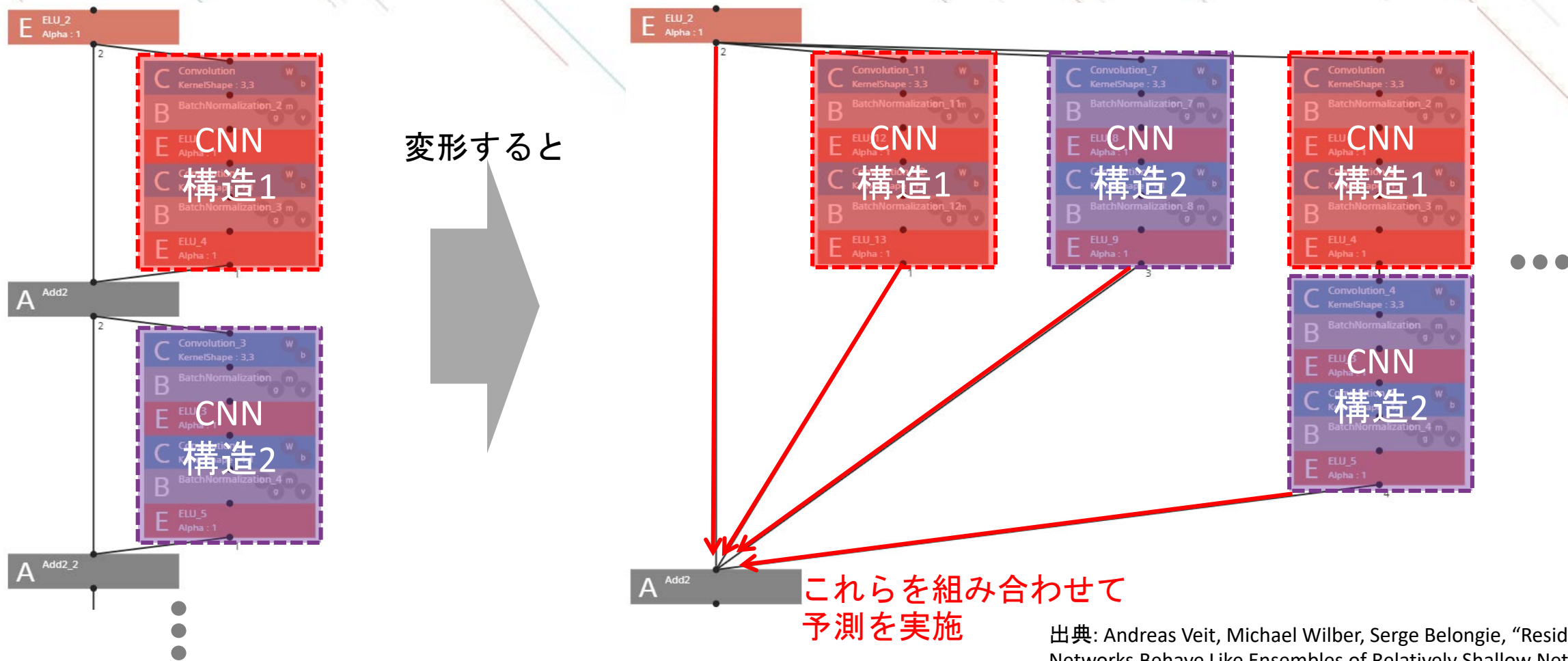
CNNの処理を通らず、そのまま次のレイヤーに繋がるショートカットコネクション

CNNの処理



(参考)Residual Network構造が高精度になる解釈

Residual Network構造はショートカットコネクションによりネットワーク内に各CNN構造を組み合わせた複数のパスが生成され、それらの組み合わせ予測を行うため、高い予測精度と汎用性を実現できると考えられています。



出典: Andreas Veit, Michael Wilber, Serge Belongie, "Residual Networks Behave Like Ensembles of Relatively Shallow Networks"



SONY

SONYはソニー株式会社の登録商標または商標です。

各ソニー製品の商品名・サービス名はソニー株式会社またはグループ各社の登録商標または商標です。その他の製品および会社名は、各社の商号、登録商標または商標です。