

Neural Network Console クラウド版

ネットワーク解説

-画像分類編-

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ResNetを用いた画像分類のサンプルプロジェクト(image_recognition.ILSVRC2012.residual_networks.resnet)のネットワーク構造を解説します。

複雑な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ので、これから画像分類を始めようという方は、画像分類の一連の流れを解説した[スターターガイド-画像分類編](#)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

本サンプルプロジェクトにはネットワークのみが含まれており、データセットは含まれておりません。ネットワークの全体像を理解いただくために、まずは想定しているデータセットなどサンプルプロジェクトの概要を解説します。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

サンプルプロジェクトの概要

2

ネットワーク構造 (共通構造)

3

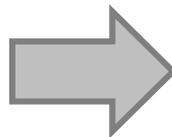
ネットワーク構造 (ネットワークタブ固有構造)

サンプルプロジェクトの概要

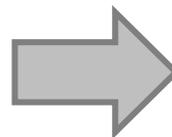
本サンプルは一般的な画像を1000個のカテゴリに分類するものです。

基礎となるネットワーク構造には ResNet(Residual Network)を利用します。ResNetは過去に画像認識コンテストで優勝した実績もあり、画像分析において高精度を実現できる構造として有名なものです。

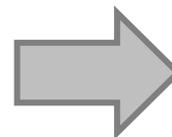
モデルの入出力



フラミンゴ



シャム猫



ダルメシアン

想定しているデータセット

本サンプルではオープンデータのImageNet^{※1}の利用を想定しています。ImageNetの画像はサイズが様々なため、本サンプル用には事前に学習用は480x480、検証用・推論用は320x320にサイズ統一する必要があります。学習用のみ高解像である理由はネットワークの前処理で生成する画像のバリエーションを増やすためです。



※1 ImageNetとはスタンフォード大学がインターネット上から収集した画像データセットで、画像検出・識別精度を競うILSVRC(ImageNet Large Scale Visual Recognition Challenge)などでも用いられる代表的なデータセットです。<http://www.image-net.org>

※2 前処理はネットワーク内で実施されます

目次

本サンプルプロジェクトにはネットワークのみが含まれており、データセットは含まれておりません。ネットワークの全体像を理解いただくために、まずは想定しているデータセットなどサンプルプロジェクトの概要を解説します。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

サンプルプロジェクトの概要

2

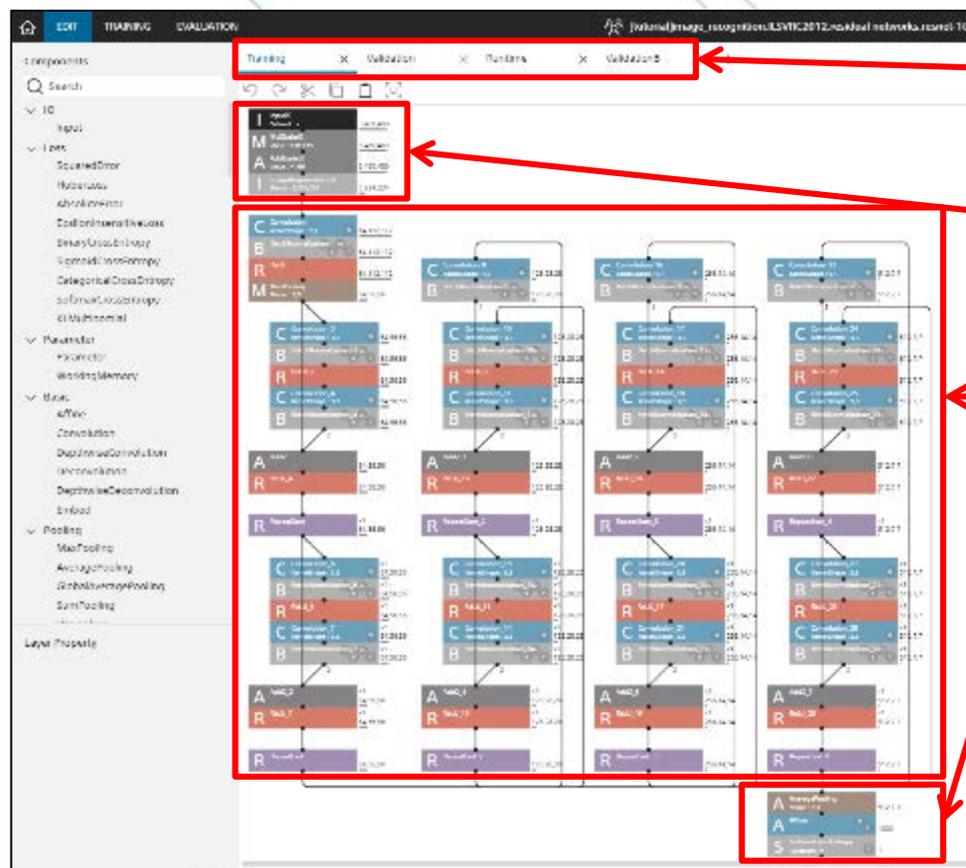
ネットワーク構造 (共通構造)

3

ネットワーク構造 (ネットワークタブ固有構造)

ネットワークの全体像

サンプルプロジェクトには学習用と推論用、2つの検証用の4つのネットワークタブがあります。各ネットワークの構成は入力部分、出力部分と全ネットワークで共通のResNet部分になります。入力部分や出力部分はネットワークタブの目的ごとに多少変化していますが、詳細は次節で説明します。



ネットワークタブ

4つのネットワークタブが存在

ネットワークの入力部分

ネットワークタブの目的ごとに構造が異なる

ResNet部分

ResNetのネットワーク構造部分で、ネットワークタブによらず共通

ネットワークの出力部分

ネットワークタブの目的ごとに構造が異なる

ResNet部分の解説

ResNetのネットワーク上の特徴は、画像分析に効果があるConvolutional Neural Network構造(CNN構造)とネットワーク処理をスキップするショートカットコネクションの2つです。次のページではこれらの概要とサンプルネットワーク上のどの部分を示しているかを説明してきます。

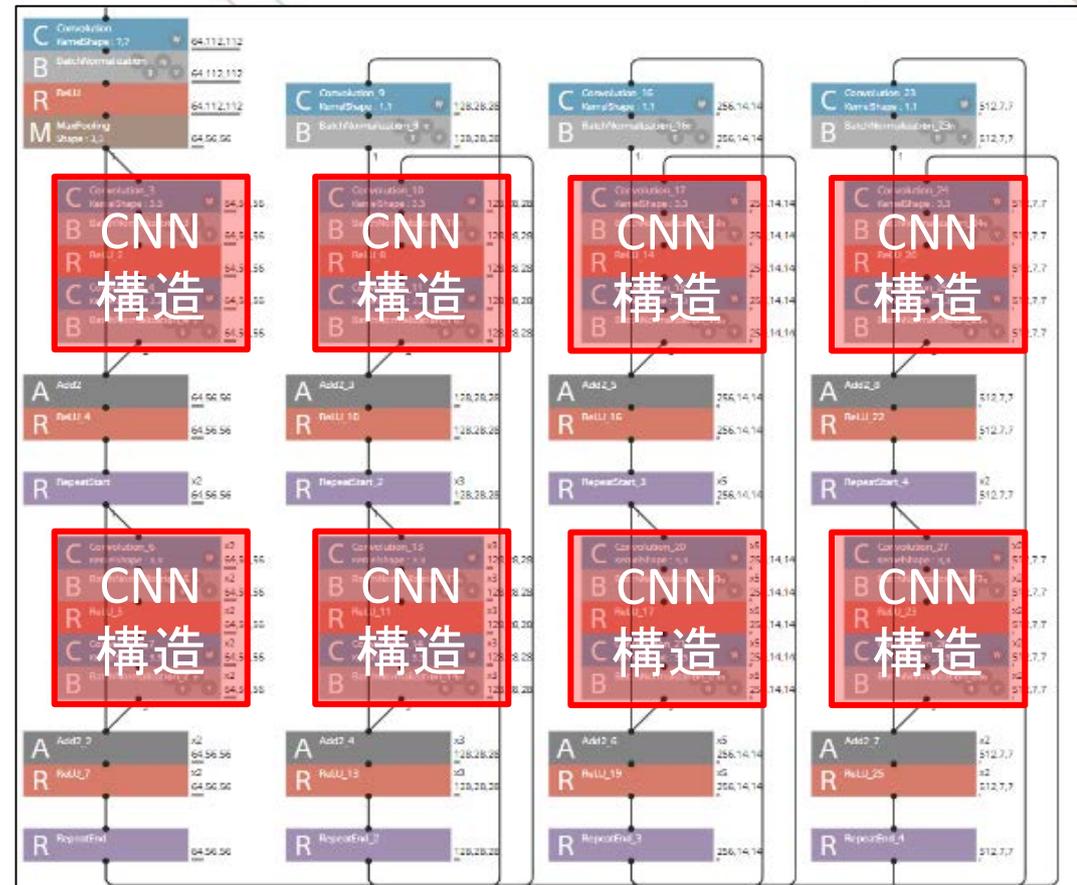
ResNetの特徴

	ネットワーク構造	概要
1	Convolutional Neural Network構造	画像分析で効果があるネットワーク構造
2	ショートカットコネクション	ネットワーク処理をスキップすることで、効率よく学習することができるネットワーク構造

- | | |
|---|--------------------------------|
| 1 | Convolutional Neural Network構造 |
| 2 | ショートカットコネクション |

Convolutional Neural Network構造

ResNetではConvolution層を用いたCNN構造が繰り返し用いられています。
CNN構造を用いることで画像の特徴量をうまく抽出することができることが知られています。
本サンプルでは次頁で説明するRepeat機能により、CNN構造の多層化を実装しています。

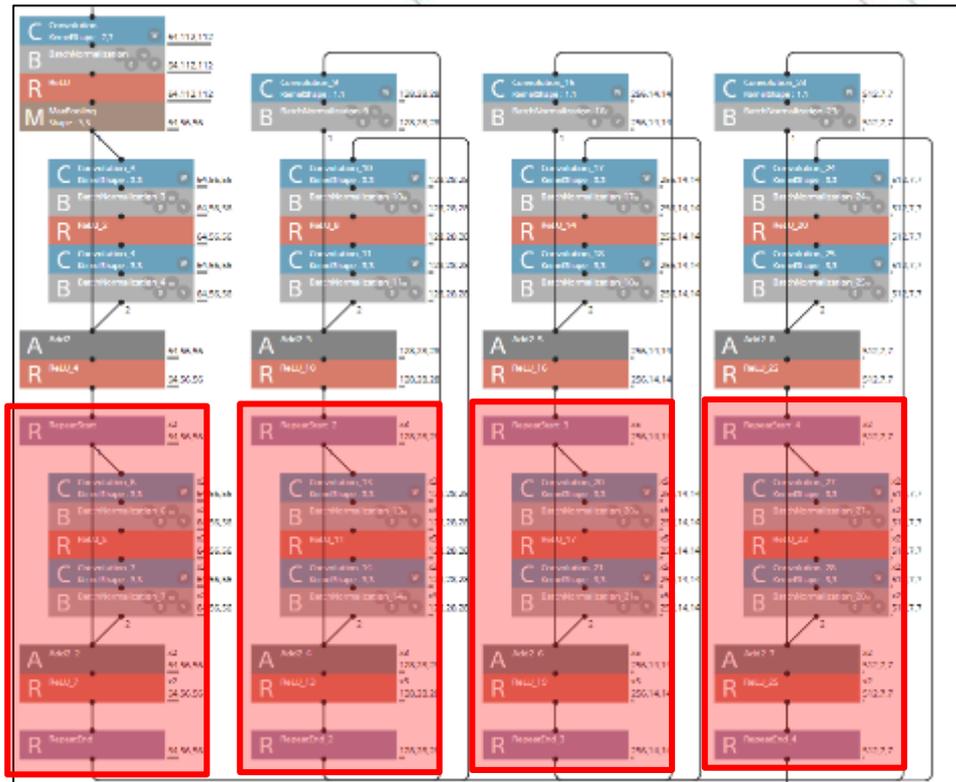


※ CNN構造の中のネットワークの組み合わせはResNetの層数によって多少変化していますが、大まかな機能は同じです。

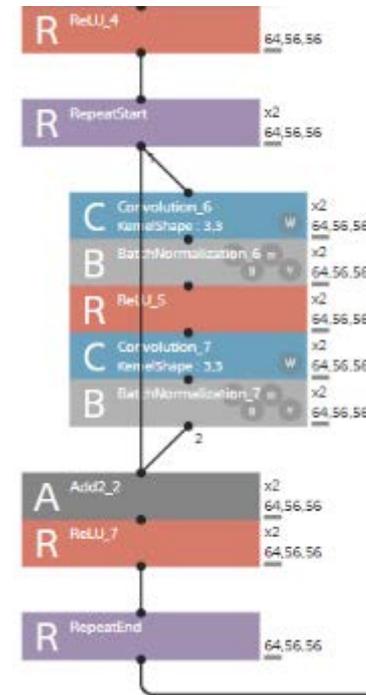
- | | |
|---|--------------------------------|
| 1 | Convolutional Neural Network構造 |
| 2 | ショートカットコネクション |

Repeat機能を用いた繰り返し

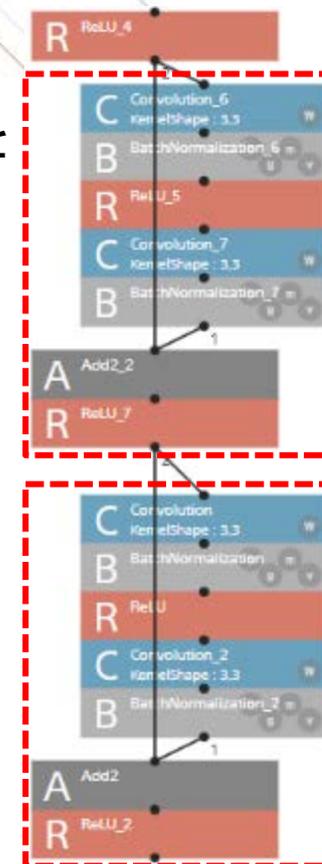
サンプルネットワークはRepeat機能(RepeatStart、RepeatEnd)を用いて、CNN構造を複数回繰り返しています。RepeatStartのオプションで繰り返し回数を設定しており、ResNetの層数が増えるとその回数が増えます。繰り返し回数が増えることでモデルが複雑になり、高精度が期待できます。



Repeatによる繰り返し



展開すると



RepeatStartからRepeatEndの中の部分が連続で繰り返し

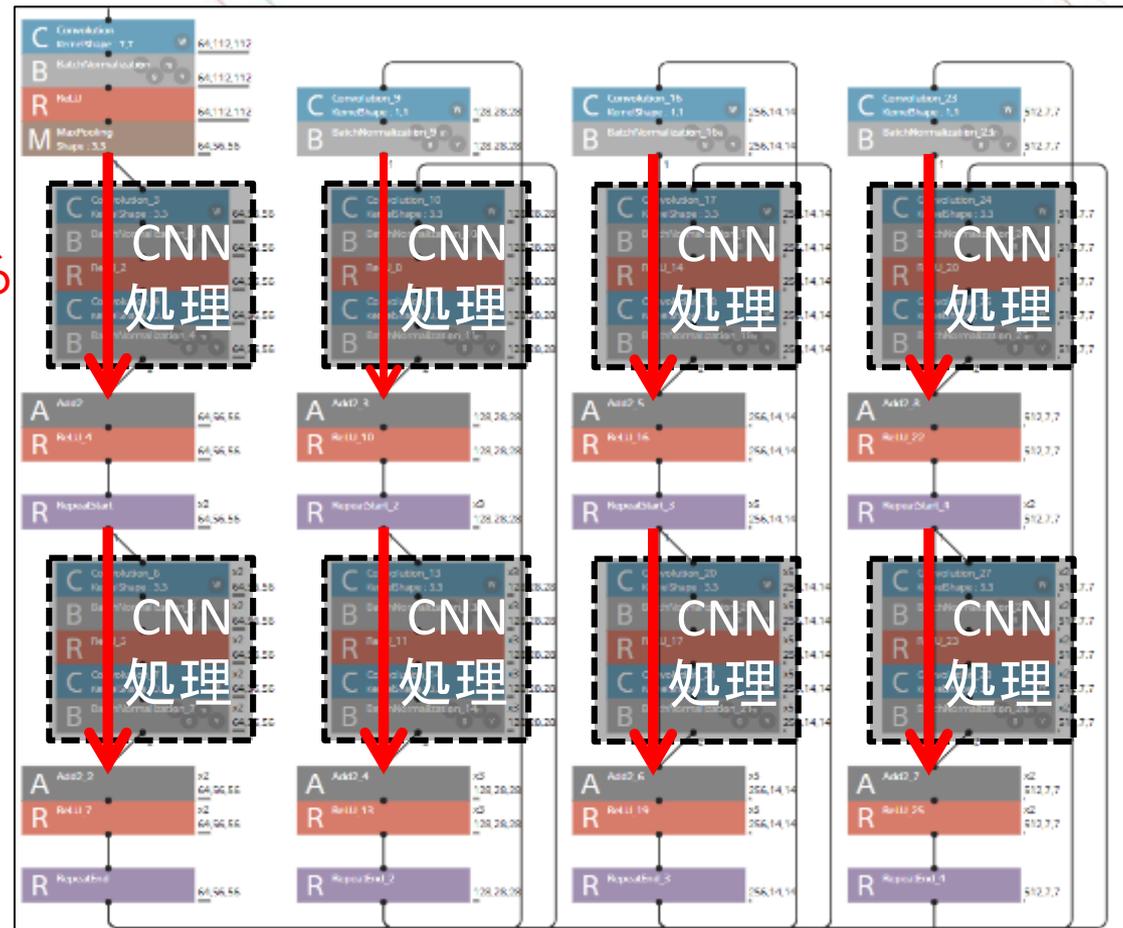
※ Repeatによる繰り返し回数はResNetの層数によって変化します。

- | | |
|---|--------------------------------|
| 1 | Convolutional Neural Network構造 |
| 2 | ショートカットコネクション |

ショートカットコネクション

ResNet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ています。これにより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([Appendix](#)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



目次

本サンプルプロジェクトにはネットワークのみが含まれており、データセットは含まれておりません。ネットワークの全体像を理解いただくために、まずは想定しているデータセットなどサンプルプロジェクトの概要を解説します。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

サンプルプロジェクトの概要

2

ネットワーク構造 (共通構造)

3

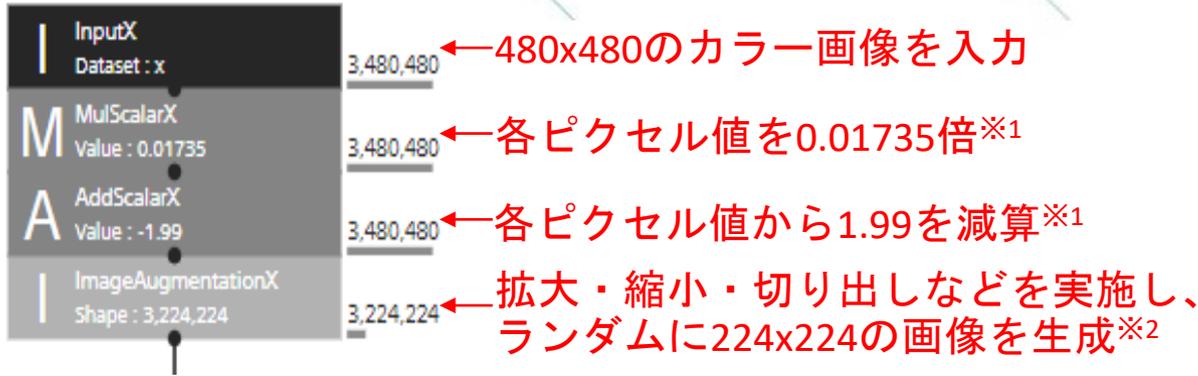
ネットワーク構造 (ネットワークタブ固有構造)

1	Training	3	Validation5
2	Validation	4	Runtime

Trainingタブの入出力部分のネットワーク

Trainingタブでは入力部分で画像のピクセル値を正規化し※1、ImageAugmentationレイヤーを用いてランダム加工し※2、ResNetに入力します。ランダム加工により1つの画像から複数のバリエーションの画像を作成することができ、学習効率を高めることができます。出力部分では学習のための誤差を算出しています。

入力部分のネットワーク



出力部分のネットワーク



※1 0.01735はImageNetの画像データの輝度の標準偏差の逆数、1.99は輝度の平均値になります。入力画像に対して標準偏差の逆数倍をし、平均値を減算することで、入力画像のピクセル値を平均0、分散1に正規化することができます。

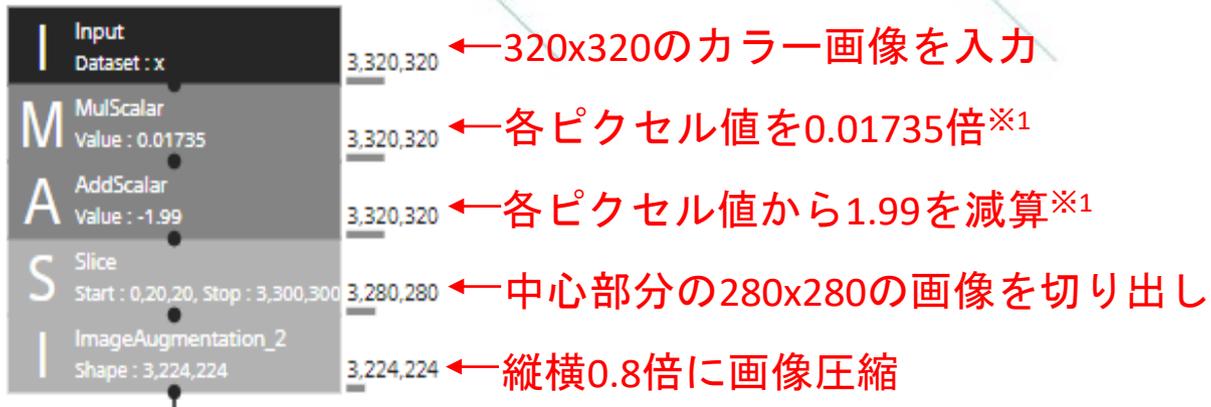
※2 ImageAugmentationによる画像生成の概要は[Appendix参照](#)

1	Training	3	Validation5
2	Validation	4	Runtime

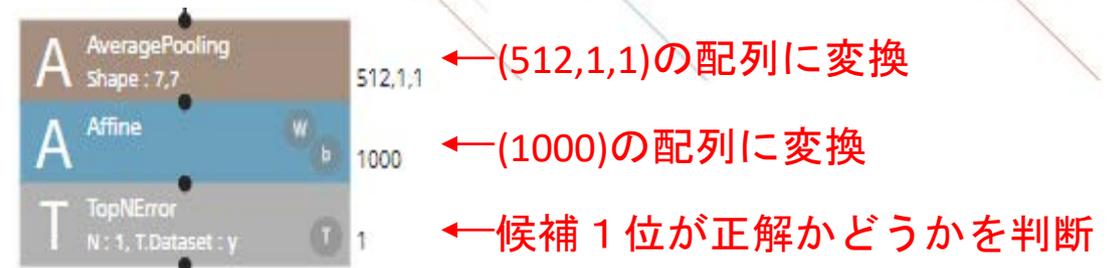
Validationタブの入出力部分のネットワーク

Validationタブでは入力部分で画像のピクセル値を正規化し※1、画像サイズを調整し、ResNetへ入力します。出力部分では1000クラスの中で候補1位のカテゴリが正解かどうかを判断します。

入力部分のネットワーク



出力部分のネットワーク



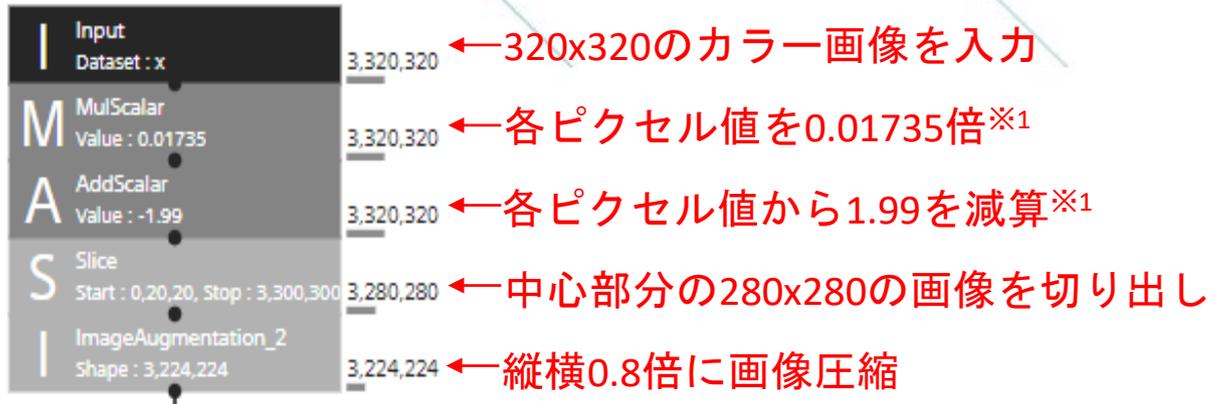
※1 0.01735はImageNetの画像データの輝度の標準偏差の逆数、1.99は輝度の平均値になります。入力画像に対して標準偏差の逆数倍をし、平均値を減算することで、入力画像のピクセル値を平均0、分散1に正規化することができます。

1	Training	3	Validation5
2	Validation	4	Runtime

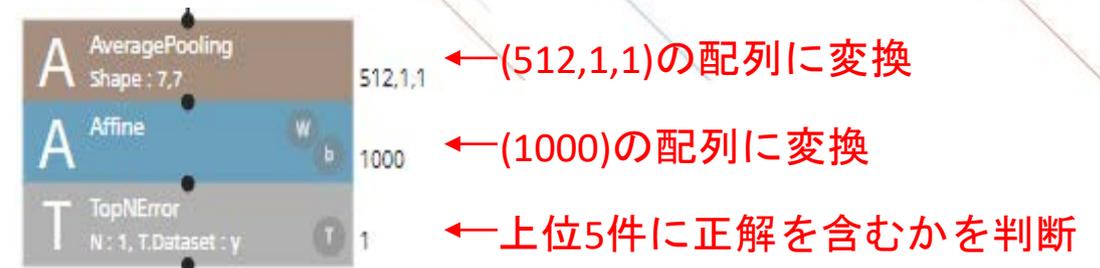
Validation5タブの入出力部分のネットワーク

Validationタブでは入力部分で画像のピクセル値を正規化し※1、画像サイズを調整し、ResNetへ入力します。出力部分では1000クラスの中で確率が高い上位5件の中に正解カテゴリが含まれるかどうかを判断します。

入力部分のネットワーク



出力部分のネットワーク



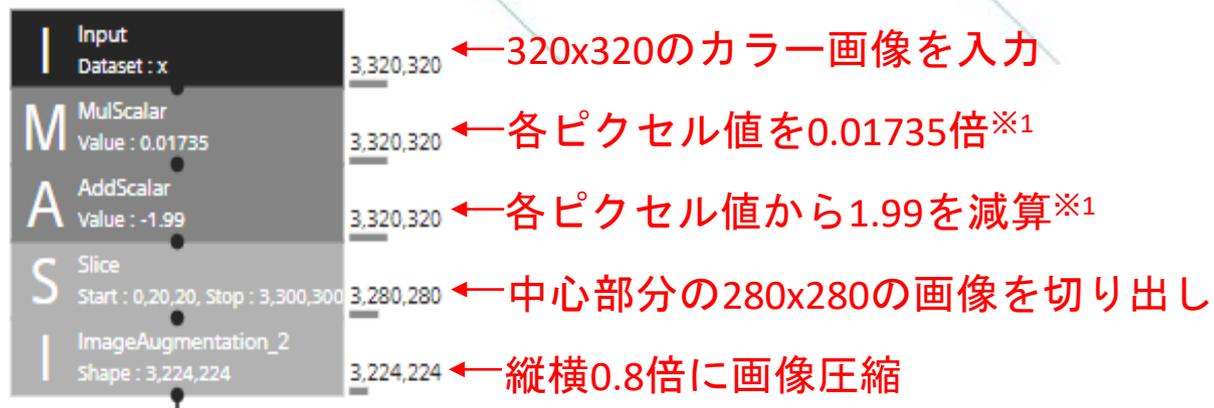
※1 0.01735はImageNetの画像データの輝度の標準偏差の逆数、1.99は輝度の平均値になります。入力画像に対して標準偏差の逆数倍をし、平均値を減算することで、入力画像のピクセル値を平均0、分散1に正規化することができます。

1	Training	3	Validation5
2	Validation	4	Runtime

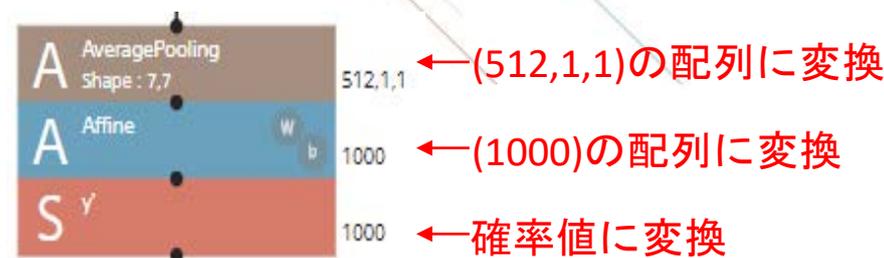
Runtimeタブの入出力部分のネットワーク

Validationタブでは入力部分で画像のピクセル値を正規化し※1、画像サイズを調整し、ResNetへ入力します。出力部分では1000クラスのそれぞれの確率値を算出しています。

入力部分のネットワーク



出力部分のネットワーク



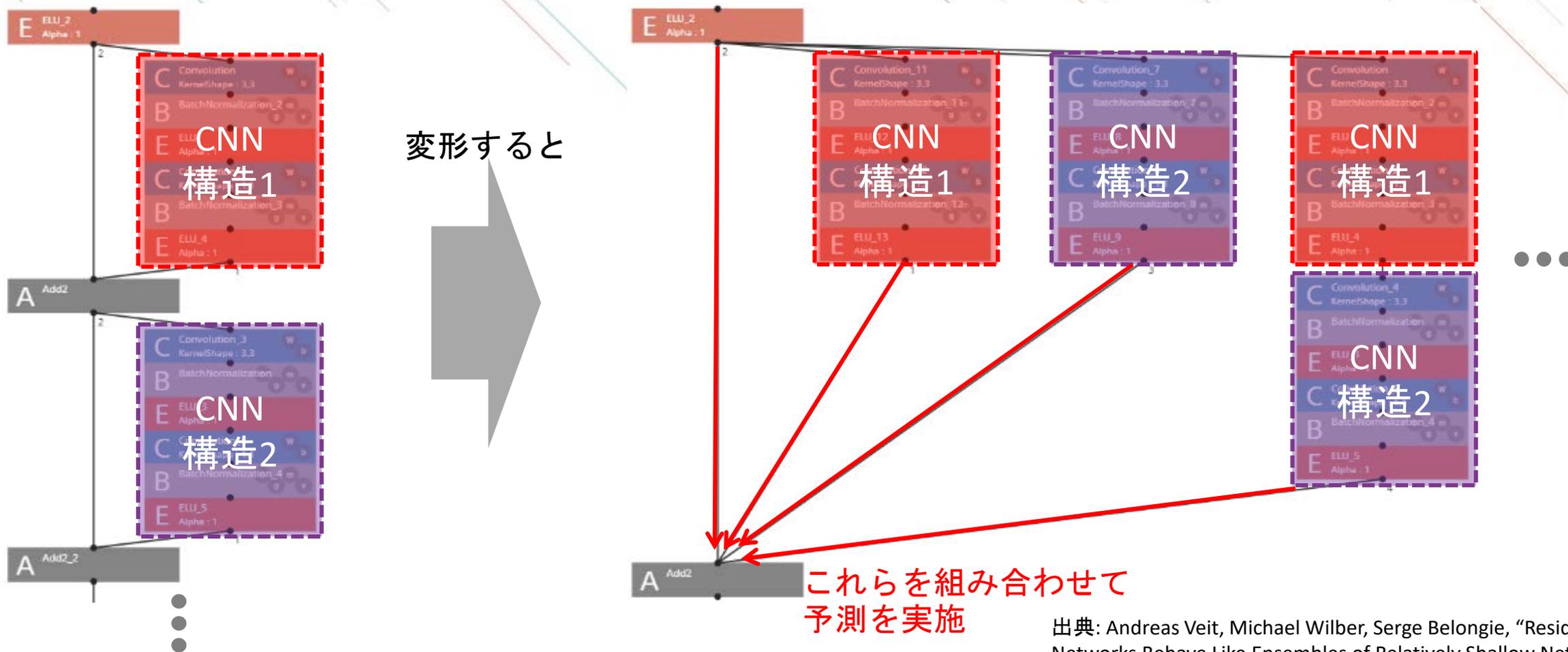
※1 0.01735はImageNetの画像データの輝度の標準偏差の逆数、1.99は輝度の平均値になります。入力画像に対して標準偏差の逆数倍をし、平均値を減算することで、入力画像のピクセル値を平均0、分散1に正規化することができます。



Appendix

(参考)ショートカットコネクションが高精度になる解釈

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"

(参考)ImageAugmentationの概要

学習時にデータの水増しを可能にするレイヤー。ランダムで切り出し、拡大縮小、回転、アスペクト比変更、歪、上下左右フリップ、輝度コントラスト、ノイズ付加が可能である。

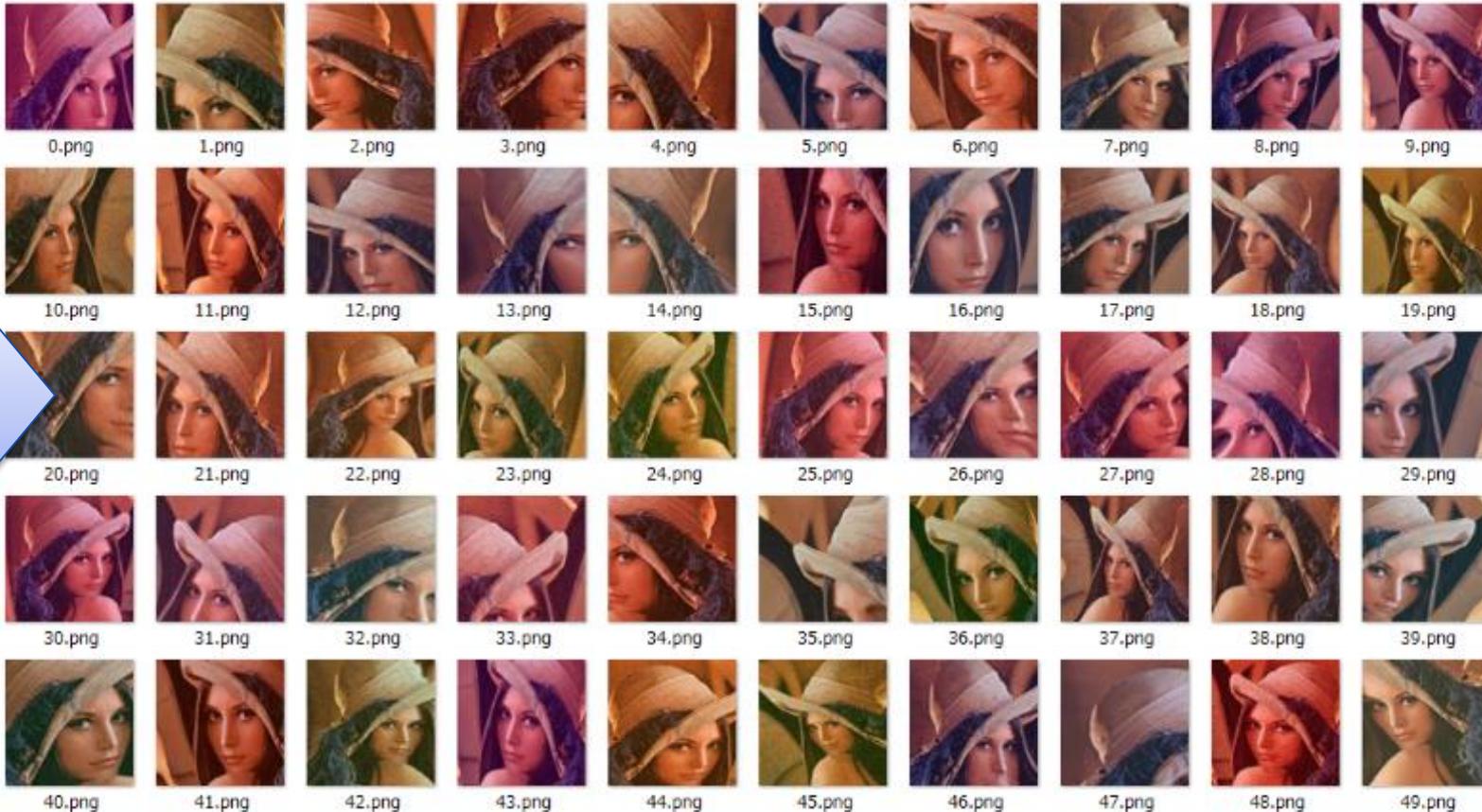
出力画像 (epoch毎にランダムに加工された画像を出力)

入力画像



Layer Property

ImageAugmentation	
Name	ImageAugmentation
Input	3, 256, 256
Shape	3,224, 224
Pad	0, 0
MinScale	0.875
MaxScale	1.5
Angle	0
AspectRatio	1.33
Distortion	0
FlipLR	True
FlipUD	False
Brightness	0
BrightnessEach	False
Contrast	1
ContrastCenter	0.5
ContrastEach	False
Noise	0
Seed	-1
SkipAtTest	False
Output	3,224, 224



0.png 1.png 2.png 3.png 4.png 5.png 6.png 7.png 8.png 9.png
10.png 11.png 12.png 13.png 14.png 15.png 16.png 17.png 18.png 19.png
20.png 21.png 22.png 23.png 24.png 25.png 26.png 27.png 28.png 29.png
30.png 31.png 32.png 33.png 34.png 35.png 36.png 37.png 38.png 39.png
40.png 41.png 42.png 43.png 44.png 45.png 46.png 47.png 48.png 49.png



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。