

3次元畳み込みニューラルネットワークを用いた タンパク質リガンド結合ポケットの予測

各務 直人^{1,2,a)} 石田 貴士^{1,2,b)}

概要: 構造ベース創薬や、構造ベースのバーチャルスクリーニングでは、標的タンパク質でリガンドが結合するポケット部位の情報が必要となる。しかしこのタンパク質リガンド結合ポケットの位置情報を得るには、X線結晶構造解析等の生物学的実験が必要であり、計算機を用いてその位置を予測する手法が求められている。既存の手法の多くは、結合ポケットの性質として与えた明示的な少数の仮定に合致する領域を探索する手法であるため、典型的な結合ポケット以外については予測精度が低下してしまう。本研究ではタンパク質リガンド結合ポケットの予測問題を3次元データにおける物体認識問題の一種と考え、3次元畳み込みニューラルネットワークを用いて予測性能の向上を図った。その結果、テストデータに対するQ2 Accuracyが0.864、ROC曲線下面積が0.932の予測精度を達成した。しかし、提案手法の性能は既存手法であるFpocketに劣っており、今後の更なる改良が必要であることが分かった。

キーワード: タンパク質リガンド結合ポケット、畳み込みニューラルネットワーク、深層学習

Prediction of Protein-Ligand Binding Sites Using 3D Convolutional Neural Networks

Abstract: In structure-based drug design and structure-based virtual screening the information of the site that a ligand binds on the target protein is needed. However, biological experiments are needed in order to obtain this protein-ligand binding site's information. Therefore, the methods predicting protein-ligand binding sites are required. Most of the existing methods search the place based on a small number of explicit hypotheses provided as typical binding site properties. Therefore, the prediction performance of proper binding sites decreases with just hypotheses. In this work, we deal protein-ligand binding site prediction problem as object recognition problem of three dimensional data, and attempted to improve the prediction performance using 3D convolutional neural networks. As a result, the Q2 accuracy for test data of 0.864 and the area under the ROC curve of 0.932 were obtained. However, the performance of the proposed method is less than that of an existing method, and it was found that it is necessary to improve.

Keywords: Protein-Ligand Binding Sites, Convolutional Neural Networks, Deep Learning

1. 導入

近年の創薬プロセスは、1個の医薬品の研究開発に約13.5年[1]の長い期間と、約26億ドル[2]という莫大な費用が必要となる。そのため、より効率的な創薬を目指し、現在では創薬のアプローチとして、計算機を用いた効率化が提案されており、既に知られている薬剤化合物(リガンド)

の知見に基づいて行われるリガンドベースの手法、そしてタンパク質の構造の情報に基づいて行われる構造ベースの手法等が用いられている。

構造ベース創薬では、タンパク質とリガンドの間の相互作用の推定等に、タンパク質リガンド結合ポケットの構造が必要となる。また現在では創薬プロセスの短縮と費用の削減を行うため、バーチャルスクリーニングという技術が用いられているが、バーチャルスクリーニングにおいて計算機により標的タンパク質に対するリガンドの活性を推定する際、タンパク質リガンド結合ポケットの位置の同定が

¹ 東京工業大学 情報理工学院 情報工学系 知能情報コース

² 東京工業大学 情報生命博士教育院

a) kakami@cb.cs.titech.ac.jp

b) ishida@cs.titech.ac.jp

不可欠である。

計算機を用いたタンパク質リガンド結合ポケットの予測問題は、タンパク質の立体構造情報を入力として、タンパク質リガンド結合ポケットの位置を予測する問題である。タンパク質は一部の例外を除いて、配列が決まれば安定する立体構造も一意に決定される。受容体であるタンパク質の立体構造にはリガンドを引き寄せ、反応しやすい位置に取り込み、反応しやすい形に変形する構造が存在する。そのため、タンパク質の立体構造情報にはタンパク質の結合ポケットの位置を決定するだけの十分な情報が存在する。タンパク質の立体構造情報はタンパク質構造データバンク (Protein Data Bank, PDB) [3] 等、様々なデータベースに集積されている。

この問題に対しては様々な手法が提案されており、主な手法には SURFNET[4], Q-SiteFinder[5], LIGSITE^{csc}[6], FINDSITE[7], 3DLigandSite[8], Fpocket[9] がある。SURFNET, Q-SiteFinder, LIGSITE^{csc} はプローブを用いてタンパク質の表面から結合ポケットの探索を行う。FINDSITE, 3DLigandSite は標的タンパク質と相溶性が高いデータベース上のタンパク質の情報をもとに結合ポケットの推定を行う。

Fpocket はアルファ球と呼ばれる概念に基づいて結合ポケットの予測を行う手法で、48 のタンパク質を含むデータセットに含まれるタンパク質の全結合ポケットの 92% 以上を識別し、現在最も高精度の予測手法の一つである。しかし、タンパク質リガンド結合ポケットの予測は、バーチャルスクリーニングにおいてタンパク質の立体構造情報から結合ポケットの情報が得られなかった場合に、リガンドの候補となる化合物と標的タンパク質との相互作用を推定する際に重要となり、さらなる予測精度の向上が期待されている。Fpocket は結合ポケットの幾何学的な性質を利用して結合ポケットの位置を予測する手法であるが、これらは結合ポケットがタンパク質のファンデルワールス表面上の溝や空洞のような場所に存在するという仮定に基づいている。すべての結合ポケットがそのような場所に存在するわけではないため、はっきりとした空洞ではない場所に位置する結合ポケットの位置の予測が困難となっている。

本研究では、深層学習の一つである 3 次元畳み込みニューラルネットワークを用いることで、タンパク質の結合ポケットの予測を行うことを目的とする。タンパク質を含む領域の、一定間隔で特定の原子を置いたときの相互作用エネルギーを複数の原子種について計算し、これらを 3 次元配列データとして 3 次元畳み込みニューラルネットワークへの入力に用いて予測を行う。

2. 提案手法

2.1 3 次元畳み込みニューラルネットワーク

機械学習の手法で、画像認識 [10]、音声認識 [11] の分野

で近年高い予測精度を示している手法に深層学習がある。深層学習は、多層のニューラルネットワークを用いて情報処理を行うことにその名が由来する手法であり、古くは 1990 年代後半から 2000 年代前半にかけて様々な研究が行われてきたが [12]、近年 Hinton らによるディープベリフネットワーク (Deep Belief Network, DBN) [13] の研究等により、画像認識や音声認識の分野で予測精度を飛躍的に向上させ、以降広く利用されている。深層学習のモデルの一種である畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) [14] や、CNN の拡張である 3 次元畳み込みニューラルネットワーク (3D Convolutional Neural Network, 3D CNN) [15] は、画像データのように 2 次元的、あるいは動画データのように 3 次元的に訓練データが並んでいる場合に予測精度を高めることが期待できる手法である。特に 3 次元畳み込みニューラルネットワークは、3 次元の物体の認識が高精度に可能なモデルである。

2.2 予測手法の概要

本研究では 3D CNN を用いた結合ポケットの予測を行う手法を提案する。本手法では、図 1 のように、対象となるタンパク質を含む領域から立方体領域を切り出し、各々の切り出した領域について、その領域に結合ポケットが含まれるかどうかを予測することで、結合ポケットの位置の予測を行う。本研究では、図 2 のように切り出した領域の相互作用エネルギーを入力とし、その領域に結合ポケットが含まれるかどうかを出力とする 2 クラス分類問題を解く。相互作用エネルギーについては 2.4 節で詳細に述べるが、数種類の 3 次元の広がりをもつデータである。CNN が扱うデータは画像認識における画像に代表されるように、主として 2 次元の広がりをもつデータであり、CNN への入力に相互作用エネルギーを用いることは一般の認識問題の自然な拡張である。CNN で提案手法のような結合ポケットの予測を行う場合、結合ポケットの性質に一定の ad hoc な仮定を置くことなく、予測に最適な結合ポケットの性質を自動的に CNN が学習でき、また CNN は小さく区切った領域での規則性を CNN が学習するので、タンパク質全体での配列相溶性の高いタンパク質が見つからないような標的タンパク質でも結合ポケットを予測できることが期待される。

2.3 データセットの構築

データセットの構築には、Astex Diverse[16] データセットを利用した。Astex Diverse データセットは、タンパク質とそれに結合するリガンドの立体構造のデータが計算機上でのドッキングに適するように選出された、多様な物性をもつ 85 個のタンパク質からなるデータセットである。Astex Diverse データセットから訓練、テストデータセッ

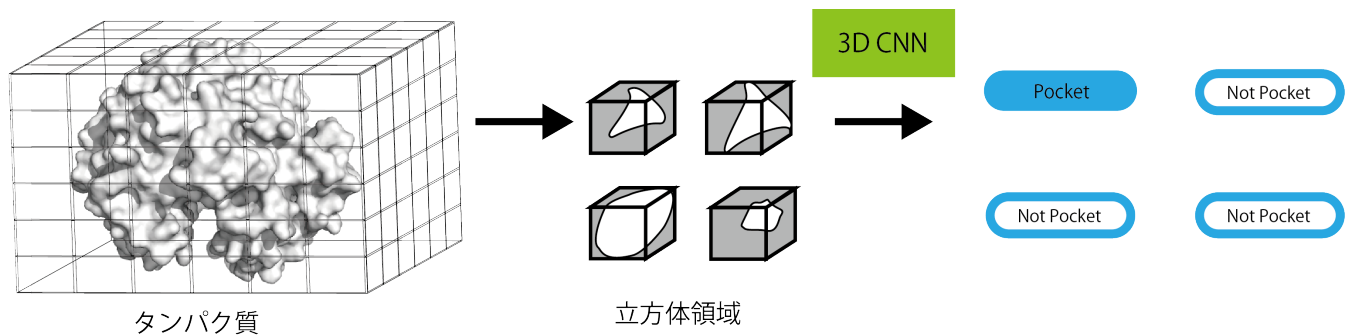


図 1 立方体領域を切り出す方法のイメージ図

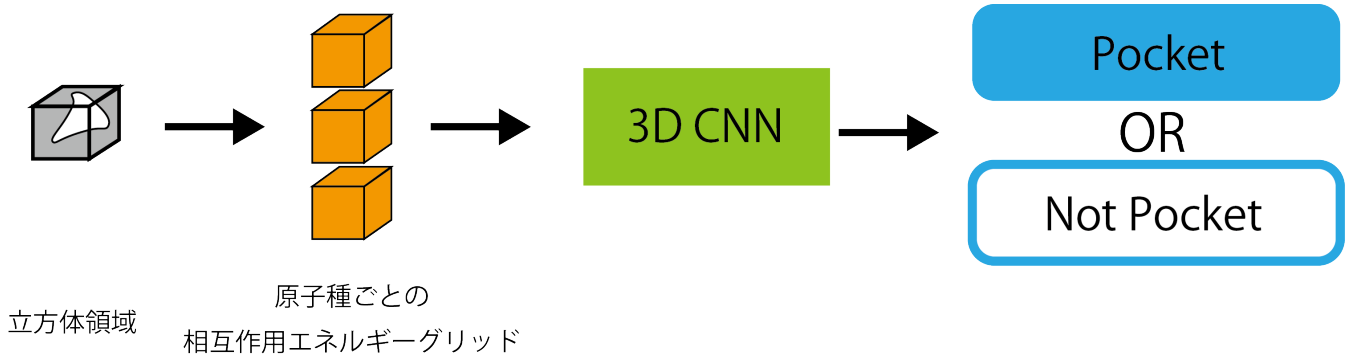


図 2 CNN への入力と出力のイメージ図

トを生成した手順は以下の通りである。

- (1) Astex Diverse データセットに含まれる 85 個のタンパク質を、5 つのサブセットに 17 個ずつランダムに振り分ける。
- (2) それぞれのタンパク質に対して、各タンパク質から立方体領域を切り出し、データの生成を行う。
- (3) 正例と負例の数を均一化させるため、各サブセットにおいて正例に対してリサンプリングを行う。
- (4) 以上の手順を経た 5 つのサブセットのうち、4 つを訓練データ、1 つをテストデータとする。

手順 2 の直後では、正例と負例の数の比率はおよそ正例 : 負例 = 1 : 742 となっている。手順 3 でリサンプリングを行うことにより、正例 : 負例 = 1 : 1 と均一化される。

また、ハイパーパラメータとして早期停止のタイミング、CNN の畳込み層のフィルタ数、CNN の学習率を決めるため、性能評価の前に交差検定を行った。訓練データとした 4 つのサブセットのうち 1 つを検証データとして、4 分割交差検定を行った。

2.4 入力特徴量

入力特徴量には、タンパク質の立方体領域内の、炭素原子、酸素原子、窒素原子のそれぞれに対するタンパク質との 3 種類の相互作用エネルギーの値を用いる。対象となる立方体領域は一辺 10 \AA であり、 $10^3 = 1000$ 個の一辺 1 \AA の小立方体に分割される。各小立方体の重心 r に炭素原子、酸素原子、窒素原子の 3 種類の原子を置いたときのタンパ

ク質との相互作用エネルギー $U(r)$ を求め、この各点の相互作用エネルギー、1 チャンネルあたり 10^3 個かつ 3 チャンネル、すなわち計 $10^3 \times 3 = 3000$ 個の値を 3 次元 CNN への入力とする。

相互作用エネルギーは AutoDock 4.0[17] を参考にし、点 r に特定の原子を置いたときのファンデルワールス相互作用エネルギー、静電相互作用エネルギー、水素結合の相互作用エネルギーの 3 つの和をとった値をその点の相互作用エネルギー $U(r)$ とした。相互作用エネルギーの計算時の定数は AutoDock 4.0 と同じものを使用している。この相互作用エネルギーは、化合物を置いたときにその化合物の構造が安定してあまりその場を離れないような場所では小さな値をとり、逆に置いたときに不安定となるような場所では大きな値をとる。タンパク質リガンド結合ポケットの付近ではリガンドの構造が安定してあまり動かなくなるため、この場所の相互作用エネルギーは小さな値をとる傾向にある。相互作用エネルギーの値が -20 kcal/mol 未満となる場合は -20 kcal/mol に、 20 kcal/mol より高くなる場合は 20 kcal/mol として扱い、CNN への入力が一定の範囲内に収まるようにクリッピングを行った。

3. 評価実験

本研究の提案手法により構築されたデータセットで訓練を行った 3D CNN による、結合ポケットの予測性能を評価するため、2 種類の評価実験を行った。一方は 3D CNN 単体の性能評価実験であり、他方は既存手法との予測性能

の比較実験である。

3.1 畳み込みニューラルネットワークの構造

本研究では4種類の構造のCNN (CNN1~CNN4)の学習を行い、交差検定を行った結果 Q2 Accuracy が最も高かったCNNを用いた。ただし、本節、3.2節で述べる交差検定と、3.3節、3.4節の性能評価に用いた2つのCNNの学習において、リガンドとみなす分子の定義等データセットの構築方法に差異がある。4種類のCNNの概要を図3に示す。CNN2~CNN4については、図3のように、中間層が畳み込み層とプーリング層を交互に2層ずつ積み重ね、その後全結合層を1層積んだ計5層になっている。ただし、CNN1についてのみ図3から2層目の畳み込み層とプーリング層(畳み込み層2とプーリング層2)をそれぞれ取り除いた構造になっている。具体的なCNNのパラメータとハイパーパラメータを表1に示す。

CNNの各層のユニットの並び方や、層の積み重ね方、学習率等は、本来ハイパーパラメータとして最適化すべきであるが、3.2節で述べるようにハイパーパラメータの探索時間を抑えるため、一部に限定して探索を行った。この学習によって得られた各CNNの、交差検定の結果得られた、訓練データと検証データのそれぞれに対する平均 Q2 Accuracy を表2に示す。この表からCNN4が最も高い Q2 Accuracy を示したため、CNN4を採用した。

3.2 ハイパーパラメータの調整

深層学習、特にCNNでは、学習時に大量のハイパーパラメータが存在し、予測性能がそれに大きく依存する。全てのハイパーパラメータに関して最適化を行おうとすると、探索範囲は組み合わせ爆発を起こし、最適化にかかる時間が非常に長くなってしまふ。よって、一部のハイパーパラメータに限定して最適化を行った。最適化しないハイパーパラメータは、一般的な深層学習における経験的に決定されることが多いハイパーパラメータ [18] を参考に、よく用いられているものを利用した。最適化したパラメータは以下の通りである。

- (1) CNN3における学習率 (10^{-3} , 10^{-5} , 10^{-7}), 全結合層のユニット数 (2^3 , 2^5 , 2^7 , 2^9)
- (2) CNN4における畳み込み層1と畳み込み層2のそれぞれのチャンネル数(畳み込み層1は 2^3 , 2^4 , 2^5 , 畳み込み層2は畳み込み層1の1倍と2倍の場合)

上のハイパーパラメータの組み合わせを変えて、5分割交差検定を行った。ハイパーパラメータの組み合わせの変え方はグリッドサーチで行った。交差検定の各foldについて、1エポックから50エポックまで学習したCNNに対して、検証データに対する平均 Q2 Accuracy を算出し、それ以上平均 Q2 Accuracy を更新しなくなったエポックで収束したとして、早期終了をした。早期終了をしたエポック

での平均 Q2 Accuracy を、ハイパーパラメータの組み合わせの間で比較し、最大となるようなハイパーパラメータの組み合わせを性能評価に用いた。

CNN3, CNN4 に対して行った交差検定の結果をそれぞれ表3, 表4に示す。最大値は強調して示している。また表4において N_1 , N_2 はそれぞれ畳み込み層1と畳み込み層2のユニット数を表す。この結果から、CNN3の全結合層のユニット数を 2^5 , 学習率を 10^{-5} とし、CNN4の畳み込み層1, 畳み込み層2のユニット数をそれぞれ 2^5 , $2^5 \times 2 = 2^6$ とした。

3.3 テストデータによる性能評価

テストデータとしたサブセットに対するCNNの Q2 Accuracy は0.864であった。また、受信者操作特性 (Receiver Operating Characteristic, ROC) 曲線とその曲線下面積 (Area Under the Curve, AUC) を図4に示す(点線の直線はランダムな出力を行う予測器のROC曲線)。今回用いた訓練データは正例と負例の数がリサンプリングにより一致しているため、予測器がランダムな出力を行う場合は、AUCの値は0.5に近い値となるが、本研究で訓練したCNNのAUCは0.932と、0.5から離れた値となっている。これらの結果から、CNNの学習自体は適切に行われており、ランダムな出力はしていないことが分かる。

3.4 既存手法との比較

本研究での提案手法の性能の評価として、テストデータに対して結合ポケットと予測された場所を上位10箇所出力させ、その中に正解がどの程度高順位に現れるかを既存手法と比較した。比較したポケット探索ソフトウェアは Fpocket である。

正解の判定方法として、結合ポケットと予測された場所の半径 8\AA 以内にタンパク質のリガンドの重心が1個あれば正解、なければ不正解として比較を行った。また、CNNの予測については、正例と予測された立方体領域の重心を、結合ポケットと予測された場所とした。

この方法により、テストデータに属する17個のタンパク質に対する予測結果を上位10箇所出力させる。予測された上位10箇所の結合ポケットに現れた正解となった結合ポケットのうち、最高順位をタンパク質の各リガンドごとに算出する。

テストデータに属する、17個のタンパク質の各リガンドに対する Fpocket と学習したCNNの予測結果を表5に示す。この表において、数値は予測された正解の最高順位であり、「-」は上位10位以内に正解が見つからなかったことを示す。タンパク質はPDB IDで表してあり、リガンド名はPDBで用いられている3文字表記で表している。本研究のCNNは、Fpocketの順位と比較すると、全23個のリガンドのうち順位が高く予測されたリガンドは1つしか

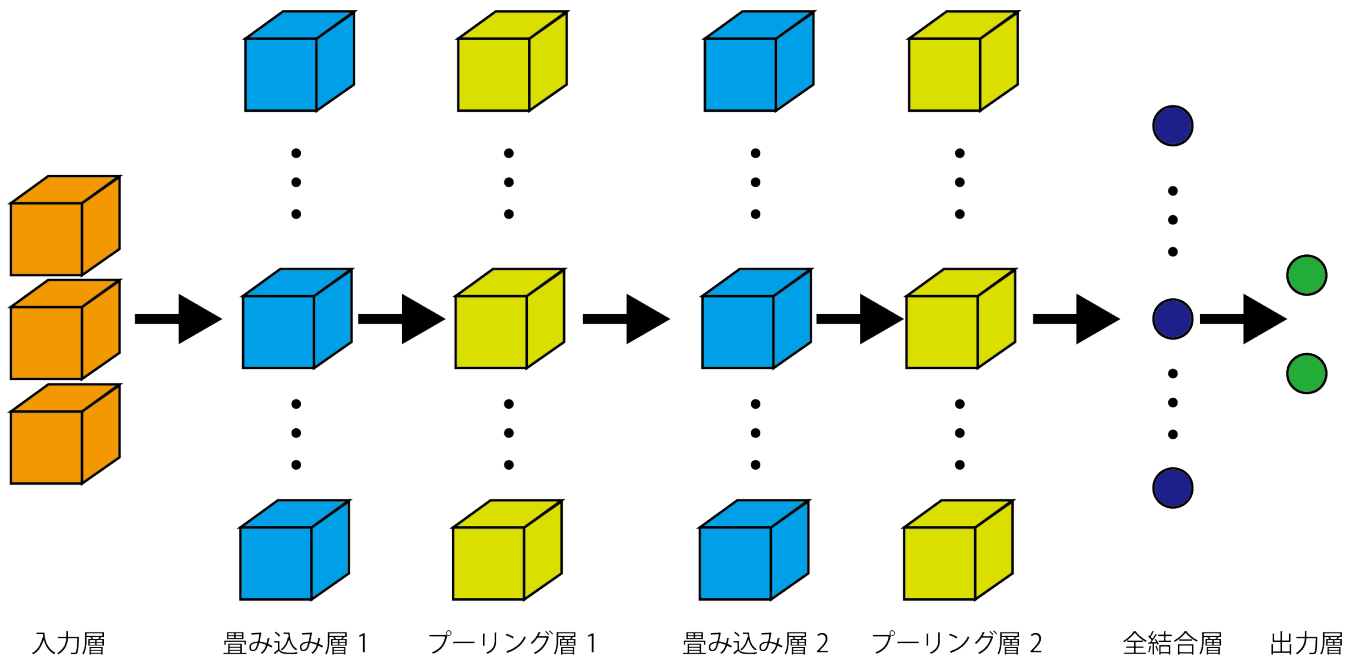


図 3 CNN の概要

表 1 3次元 CNN による訓練のパラメータ・ハイパーパラメータ

		CNN1	CNN2	CNN3	CNN4
ミニバッチのサイズ		100			
モメンタムの係数		0.9			
L2 正則化の係数		4×10^{-3}			
学習率		10^{-5}	10^{-5}		
コスト関数		交差エントロピー関数			
畳み込み層 1	チャンネル数	96			32
	活性化関数	ReLU			
	フィルタサイズ	$3 \times 3 \times 3$			
	パディング	1			
	ストライド	1			
プーリング層 1	プーリング関数	最大プーリング			
	プーリングサイズ	$3 \times 3 \times 3$			
	ストライド	2	1		
畳み込み層 2	チャンネル数	(なし)	256	64	
	活性化関数	(なし)	ReLU		
	フィルタサイズ	(なし)	$3 \times 3 \times 3$		
	パディング	(なし)	1		
	ストライド	(なし)	1		
プーリング層 2	プーリング関数	(なし)	最大プーリング		
	プーリングサイズ	(なし)	$3 \times 3 \times 3$		
	ストライド	(なし)	2	1	
全結合層	ユニット数	48			8
	活性化関数	ReLU			
出力層	ユニット数	2			
	活性化関数	ソフトマックス関数			

く、性能が大きく劣っていた。

4. 考察

本研究ではハイパーパラメータの探索時間を削減するため、一部のハイパーパラメータに限定して調整を行った。

この調整が不足していたために、CNN が性能を発揮できなかった可能性が存在する。

また、本研究で学習を行った CNN はいずれも学習率を、学習のどの段階でも、構造のどの場所でも等しい値に設定したが、学習の進み具合に応じて学習率を変化させる、学

表 2 交差検定によって得られた 4 種類の CNN の、学習データと検証データのそれぞれに対する平均 Q2 Accuracy

	CNN1	CNN2	CNN3	CNN4
訓練データ	0.630	0.740	0.698	0.804
検証データ	0.592	0.696	0.693	0.702

表 3 グリッドサーチにおける各ハイパーパラメータの組み合わせに対する CNN3 の検証データの Q2 Accuracy

ユニット数 \ 学習率	10^{-3}	10^{-5}	10^{-7}
2^3	0.510	0.692	0.637
2^5	0.501	0.693	0.627
2^7	0.501	0.639	0.628
2^9	0.506	0.638	0.633

表 4 グリッドサーチにおける各ハイパーパラメータの組み合わせに対する CNN4 の検証データの Q2 Accuracy

$N_2 \setminus N_1$	2^3	2^4	2^5
N_1	0.6882	0.6927	0.7016
$2N_1$	0.6873	0.6964	0.7017

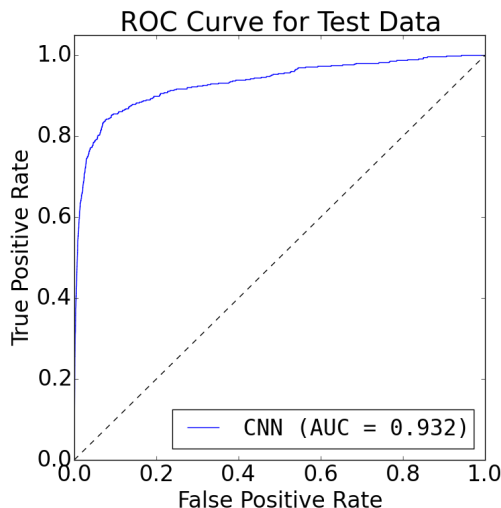


図 4 CNN のテストデータに対する ROC 曲線と AUC の値

習の進みやすい場所や進みにくい場所での学習率を変化させる等すれば、CNN がより高い性能を発揮できた可能性がある。

この他、本研究では予測されたタンパク質に対する上位 10 個を出力している。この 10 個という個数はタンパク質 1 個に対する立方体領域の数に対して非常に少ない。また False Positive のデータが多いため、True Positive のデータが上位 10 個に現れにくくなっている。このことが Fpocket との性能比較時に性能が劣っていることを助長している可能性がある。

5. 結論

本研究では多くの認識問題 [10], [11] に対して高い認識精度を示している 3 次元畳み込みニューラルネットワークを用いて、タンパク質リガンド結合ポケットの予測を行った。

表 5 テストデータに属する 17 個のタンパク質に対する予測結果

PDB ID	分子名	主鎖名	CNN	Fpocket
1g9v	RQ3	A	-	1
	RQ3	C	-	-
1ia1	TQ3	A	4	-
	TQ3	B	-	3
1k3u	IAD	A	-	2
1ke5	LS1	A	-	1
1p2y	NCT	A	-	1
1p62	GEO	B	1	7
1r55	097	A	-	1
1r9o	FLP	A	-	1
	NDR	A	-	2
1sqn	NDR	B	-	4
	RRC	A	-	1
1uou	CMU	A	3	2
1w1p	GIO	A	-	-
	GIO	B	-	5
1w2g	THM	A	-	1
	THM	B	-	2
1xoq	ROF	A	-	1
	ROF	B	-	2
1y6b	AAX	A	-	2
1ywr	LI9	A	-	1
1z95	198	A	-	1

タンパク質リガンド結合ポケットの予測を行う既存の手法としては Fpocket[9] があったが、十分な信頼がおける程度の予測精度をもっていなかった。多くの認識問題で高い認識精度を示している深層学習の手法を適用することで、結合ポケットの予測精度の向上させることができる可能性があった。そこで本研究では深層学習のモデルの一種である 3 次元畳み込みニューラルネットワークを利用してタンパク質リガンド結合ポケットの予測を行った。

本研究での提案手法の予測精度の評価を行うため、既存手法の 1 つである Fpocket との比較実験を行った。訓練データ、テストデータには Astex Diverse データセットを用いて CNN の学習を行い、テストデータに属する各タンパク質の各リガンドに対する正解の平均順位で性能の比較を行った。しかし Fpocket に予測精度が劣る結果となった。本研究では調整にかかる時間の削減のため一部に限定して行った、ハイパーパラメータの調整の不足は予測精度が劣った原因の一つであると考えられる。また、False Positive のデータが多いため予測した上位 10 個の立方体領域に True Positive が現れにくくなっていることも予測精度を下げている要因と考えられる。

また、ハイパーパラメータの調整の不足という問題の解消については、ランダムサーチ等の効率的な手法も取り入れた、より十分な範囲での調整が必要である。学習率を固定させたことに関しては、ある程度学習が進んだ時点で学習率を 10 分の 1 にし、場合によってはこれを数回繰り返

す手法 [19] や、頻出する誤差勾配の成分よりもまれに出現する誤差勾配の成分をより重視して学習するように学習率を自動的に決定する手法 [20] がより良く学習できることが示されており、これらの手法を用いれば、CNN の学習がより効率的に進むことが期待できる。

True Positive のデータが上位 10 個に現れにくくなっている問題に対しては、正例と判定された立方体領域同士で距離的に近い立方体領域はクラスタリングして 1 つにまとめる方法が有効である。これにより正例と予測される立方体領域の実質的な総数が減り、True Positive のデータが上位 10 個に現れる可能性が高まる。

謝辞 本研究は JSPS 科研費 15K16082 の助成を受けたものである。

参考文献

- [1] Paul S. M., Mytelka D. S., Dunwiddie C. T., How to improve R&D productivity: the pharmaceutical industry's grand challenge, *Nature Reviews Drug Discovery*, 9, 203–214, 2010.
- [2] Mullard A., New drugs cost US\$2.6 billion to develop, *Nature Reviews Drug Discovery*, 13, 877, 2014.
- [3] RCSB Protein Data Bank.
<http://www.rcsb.org/pdb/> (2016/6/1 閲覧).
- [4] Laskowski R. A., SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions, *Journal of Molecular Graphics*, 13(5), 323–330, 307–308, 1995.
- [5] Laurie A. T., Jackson R. M., Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, 21, 1908–1916, 2005.
- [6] Huang B., Schroeder M., LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation, *BMC Structural Biology*, 6(19), 2006.
- [7] Skolnick J., Brylinski M., FINDSITE: a combined evolution/structure-based approach to protein function prediction. *Briefings in Bioinformatics*, 10, 378–391, 2009.
- [8] Wass M. N., Kelley L. A., Sternberg M. J., 3DLigand-Site: predicting ligand-binding sites using similar structures. *Nucleic Acids Research*, 38, W469–W473, 2010.
- [9] Guilloux V. L., Schmidtke P., Tuffery P., Fpocket: An open source platform for ligand pocket detection, *BMC Bioinformatics*, 10, 168, 2009.
- [10] Olga R., Jia D., Hao S., Jonathan K., Sanjeev S., Sean M., Zhiheng H., Andrej K., Aditya K., Michael B., Alexander C. B. and Li F., ImageNet Large Scale Visual Recognition Challenge, *arXiv:1409.0575*, 2014.
- [11] Hinton G., Deng L., Yu D., Dahl G., Mohamed A., Jaitly N., Senior A., Vanhoucke V., Nguyen P., Sainath T. and Kingsbury B., Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, *IEEE Signal Processing Magazine*, 29, 82–97, 2012.
- [12] Rumelhart D. E., McClelland J., Parallel Distributed Processing: Explorations in the Microstructure of Cognition, *MIT Press*, 1986.
- [13] Hinton G., Osindero S. and Teh Y. W., A Fast Learning Algorithm for Deep Belief Nets., *Neural Computation*, 18, 1527–1554, 2006.
- [14] LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D., Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, 1(4), 541–551, 1989.
- [15] Ji S., Xu W., Yang M., Yu K., 3D Convolutional Neural Networks for Human Action Recognition, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(1), 221–231, 2013.
- [16] Hartshorn M. J., Verdonk M. L., Chessari G., Brewerton S. C., Mooij W., Mortenson P. N., Murray C. W., Diverse, High-Quality Test Set for the Validation of Protein - Ligand Docking Performance, *Journal of Medicinal Chemistry*, 50(4), 726–741, 2007.
- [17] Morris G. M., Goodsell D. S., Huey R., Olson A. J., Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4, 10(4), 293–304, 1996.
- [18] Bengio Y., Practical recommendations for gradient-based training of deep architectures. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7700 LECTU, 437–478, 2012.
- [19] 岡谷貴之, (2015), 機械学習プロフェッショナルシリーズ 深層学習, 講談社.
- [20] Duchi J., Hazan E., Singer Y., Adaptive Subgradient Methods for Online Learning and Stochastic Optimization, *Journal of Machine Learning Research*, 12, 2121–2159, 2011.