

食べ物の好みの脳計算過程：深層学習によるモデリング

Neural computations underlying food preferences: modeling with deep learning.

青木 悠飛^{*1}

Yuhi Aoki

遠山 朝子^{*1}

Asako Toyama

山下 祐一^{*2}

Yuichi Yamashita

鈴木 真介^{*1}

Shinsuke Suzuki

^{*1}一橋大学

Hitotsubashi University

^{*2}国立精神・神経医療研究センター

National Center of Neurology and Psychiatry

Our daily dietary choices are guided by the subjective values we assign to foods. Yet, little is known about how the subjective values are constructed in the brain. The present study aims to elucidate the neurocomputational processes underlying the food valuation using a deep convolutional neural network model (DCNN). Notably, DCNNs perform at human-level accuracy in object recognition tasks by processing visual information in a manner similar to the brain, making them a suitable computational model for food valuation based on visual inputs. By analyzing rating data for 896 food images provided by 200 participants, we demonstrated that DCNNs can significantly predict the subjective values of foods based on their visual features. Furthermore, an examination of neural activity patterns across the layers of the DCNN revealed that higher-order attributes such as subjective value, healthiness, and tastiness are represented in the later layers, while low-level visual information (i.e., color) is consistently encoded across both early and late layers. These findings suggest that low-level visual information plays a critical role in the entire process of subjective value computation in the brain. Future research will compare DCNN predictions with neuroimaging data to deepen our understanding of the neural computations involved in food valuation.

1. はじめに

摂食行動はヒトの生存に不可欠であり、日常生活において「何を食べるか」という選択は重要な関心事の一つである。このような意思決定は、各選択肢の主観的価値を比較することで行われると考えられている [Rangel 08]。食品に対する偏った嗜好は肥満や摂食障害と関連することが知られており [Foerde 15, Spinelli 21]、食品の主観的価値がどのように計算されるのかを明らかにすることは、ヒトのウェルビーイングの向上に資すると思われる。

ヒトの脳内において、主観的価値はどのように計算されているのか？食品の主観的価値は、栄養価、味、健康への影響、見た目、価格などの様々な要因に影響されることが知られている [Hare 09, Suzuki 17, Motoki 20]。また、最終的に計算された主観的価値の情報は腹内側前頭前野で表象されていることが明らかになっている [Chib 09]。しかしながら、主観的価値がどのような過程を経て計算されるのかについては未解明である。

近年、深層畳み込みニューラルネットワーク (DCNN) とヒトの行動や脳活動を対比させることで、脳における知覚、認知、価値判断の仕組みを探る試みが行われている [Yamins 14, Igaya 21, Igaya 23]。DCNN は脳の階層的な視覚情報処理機構を模しており、画像分類においてヒトと同等の性能を発揮することから、脳の計算論モデルとして適している。例えば、Igay et al. は絵画の主観的価値を予測する DCNN を構築し、その性質を解析した。その結果、絵画の低次の画像情報は DCNN の低次層に、高次の情報 (抽象的属性など) は高次層に保持されていることを見出した [Igay 21]。

本研究では、食品画像に対する主観的価値 (好み) の計算過程を、DCNN によるモデル化を用いて明らかにすることを目的とする。具体的には、食品画像の主観的価値を予測する DCNN を構築し、各層の活性化パターンを解析することで、

連絡先: 青木 悠飛, 一橋大学 ソーシャル・データサイエンス研究科, dm240001@g.hit-u.ac.jp, <https://github.com/edegp>

様々な情報 (低次の画像情報、栄養価、味、健康への影響) の表現を検証する。これにより、脳が食品の画像情報をどのように処理・統合し、最終的な主観的価値へと変換しているのかを理解することを目指す。

2. 方法

2.1 オンライン調査

被験者 200 名 (年齢: 平均 39.08 歳 [20–72 歳]) が Food-Pics [Blechert 14, Blechert 19] に含まれる 896 枚の食品画像について、「好きか」 (主観的価値)、「美味しそうか」、「健康的か」を評定した。評定には、1 (まったく同意しない) から 8 (強く同意する) までの 8 段階リッカート尺度を用いた。データ照合の不備により除外された 1 名を除く 199 名分の評価の平均を各画像の評定値とした (主観的価値の上位と下位の画像は図 1 参照)。調査は国立精神・神経医療研究センターの倫理審査委員会の承認を得て行われた。



図 1: 食品画像の主観的価値上位 5 個と下位 5 個

2.2 DCNN モデル

食品画像から主観的価値を推定するため、代表的な DCNN である ConvNeXt [Liu 22]、ResNet [He 15]、VGG [Simonyan 15] を用いた。いずれも ImageNet-1k で事前学習されており、最終の全結合層の出力が主観的価値 (1 から 8

の連続値)となるように付け替え、すべての層のパラメータをファインチューニングした。学習データは、水増し処理を5回繰り返した食品画像と元の画像を使用した。水増しの具体的な処理は以下のとおりである：リサイズ (240×240)、センタークロップ (224×224)、水平反転、アフィン変換 (回転±20°、移動上下左右20%、スケール70~120%)、ガウシアンフィルタ (カーネルサイズ5×5, $\sigma = 0.01 \sim 4.0$)、色変換 (明るさ7.5%、色相3%、彩度3%)、正規化 (学習画像の平均と標準偏差に基づく)。損失関数には Huber 損失 ($\delta = 1$) を使い、最適化手法は AdamW とした。学習率は 0.0001、ミニバッチサイズは 373 (GPU メモリの上限)、エポック数は 250 に設定した。評価には 6 分割交差検証を採用し、各分割の検証データで得られた予測値を結合して実測値との相関係数を算出した。相関係数の統計的有意性は並べ替え検定 (並べ替え回数 = 50) により検証した。環境は Intel Xeon w5-2465X processors と RTX 4000 Ada を用いて Python 3.12.7、PyTorch 2.5.0 で機器は CUDA を使用した。

2.3 DCNN 各層のデコーディング解析

学習済み DCNN の各層の活性化パターンを解析することで、食品画像に含まれる様々な情報 (低次の画像情報、栄養価、味、健康への影響) がどの層で表現されているのかを検証した。解析にはリッジ回帰を用い、各層の活性化パターンから画像情報をどの程度説明できるのかを評価した。

まず、各層ごとに、896 枚の食品画像に対する各ニューロンの活動 (畳み込み層の活性化関数の出力) を結合した $896 \times d$ 次元の活性化パターン行列を得る (d は当該層のニューロン数)。次に、活性化パターン行列を主成分分析で累積寄与率 80% に達するまで次元削減し、その主成分を説明変数として Ridge 回帰により画像の各種属性 (主観的価値、美味しさ、健康的か、色情報、栄養価) を推定した。

正則化パラメータは 8 分割交差検証により最適化し、どの層がどの属性を表現しているかを評価した。説明力の指標には推定値と実測値の相関係数を使用した。

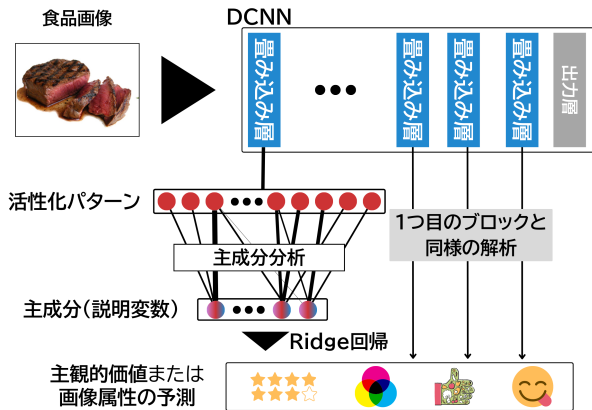


図 2: DCNN 各層のデコーディング解析のイメージ。

3. 結果

3.1 DCNN の予測精度

DCNN は食品画像の主観的価値を予測できた (図 3)。交差検証で評価した予測精度 (予測値と実測値の相関係数) は ConvNext で 0.680 ($p < 0.02$)、ResNet で 0.608 ($p < 0.02$)、VGG16 で 0.577 ($p < 0.02$) となった。いずれのモデルでも主観的価値を統計的有意に予測できたが、とくに ConvNeXt の精度が高かった。

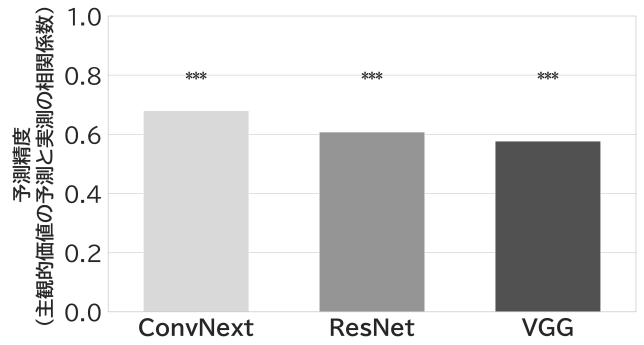


図 3: 各モデルの予測精度 (予測値と実測値の相関係数)

3.2 DCNN 各層の情報表現

次に、最も精度の高かった ConvNeXt の各層の活性化パターンを解析した (図 4)。主観的価値や美味しさ、健康への影響といった高次情報は、低次層の活性化パターンからは高い説明力を示さなかったが、高次層になるほど説明力が上がった。一方、色 (赤・緑・青) のような低次情報は、低次層の活性化パターンから説明力が高いだけでなく、高次層でも高い説明力があった。また、栄養価は全般的に説明力が低かった。

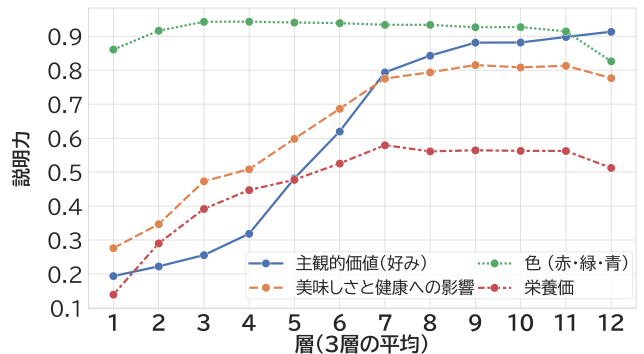


図 4: ConvNext の各層の活性化パターンによる各画像の属性の説明力 (推定値と実測値の相関係数の3層ごとの平均値)

4. 考察

本研究では、DCNN を用いることで、食品画像の主観的価値を予測できることを示した。さらに、各層の活性化パターンの分析の結果、美味しさ、健康への影響、主観的価値といった高次情報は主に高次層で表現されている一方、低次の色情報は低次層から高次層まで広く表現されていることが明らかとなった。また、栄養価に関しては、顕著な表現は見られなかった。これらの結果は、低次の画像情報が食品の主観的価値の計算過程全体を通じて重要な影響を及ぼしていることを示唆する。

上記の情報表現は絵画の主観的価値に関する先行研究 [Iigaya 21] とは異なっている。絵画の主観的価値を予測する DCNN においては、低次の色情報は低次層で主に表現されており、高次層では表現されていないことが示されている。本研究の結果から、食品画像の評価では絵画の場合とは異なり、色情報が主観的価値の推定において重要な役割を果たし、その計算過程全体を通して保持される可能性が示唆された。つまり、色情報が段階処理において抽象化されず維持されるという可能性を示唆する。

また、栄養価に関する情報は、DCNN の活性化パターンにおいて十分に表現されていなかった。この結果は、主観的栄養価が主観的価値の予測に有効であることを示した先行研究 [Suzuki 17] とは一致しない。栄養価が主観的価値の計算にお

いてどのような役割を果たすのかを検証することは、今後の課題である。

本研究では、被験者全体の平均的な主観的価値を用いて解析を行ったが、肥満群や特定の食習慣を有する集団への適用可能性については今後の課題である。先行研究では、肥満者や摂食障害患者は健常者とは異なる食嗜好を示すことが報告されている [Foerde 15, Spinelli 21]。このことから、健常者と肥満者では、視覚的特徴の重み付けや主観的価値の計算過程が異なる可能性がある。

最後に、DCNN は視覚情報を階層的に学習するモデルであり、食品の意味や概念に関する情報を直接的に処理するメカニズムを持たない。また、視覚情報の処理においても、DCNN はヒトとは異なる特性を持つことが、最近の研究で明らかになっている [Bowers 23, Caplette 24]。したがって、ヒトの主観的価値の計算過程を包括的に理解するためには、食品に関する意味情報や概念情報を統合したマルチモーダルモデルの開発が不可欠である。さらに、視覚情報以外の要因を含めた包括的なモデルによる検証も、今後の課題として残されている。加えて、DCNN と実際の脳活動データ（機能的 MRI や脳波など）との対応関係を検証することも、今後の重要な研究課題となる。

参考文献

- [Blechert 14] Blechert, J., Meule, A., Busch, N. A., and Ohla, K.: Food-pics: an image database for experimental research on eating and appetite, *Frontiers in Psychology*, Vol. 5, (2014)
- [Blechert 19] Blechert, J., Lender, A., Polk, S., Busch, N. A., and Ohla, K.: Food-Pics_Extended—An Image Database for Experimental Research on Eating and Appetite: Additional Images, Normative Ratings and an Updated Review, *Frontiers in Psychology*, Vol. 10, (2019)
- [Bowers 23] Bowers, J. S., Malhotra, G., Dujmović, M., Llera Montero, M., Tsvetkov, C., Biscione, V., Puebla, G., Adolphi, F., Hummel, J. E., Heaton, R. F., and al., et : Deep problems with neural network models of human vision, *Behavioral and Brain Sciences*, Vol. 46, p. e385 (2023)
- [Caplette 24] Caplette, L. and Turk-Browne, N. B.: Computational reconstruction of mental representations using human behavior, *Nature Communications*, Vol. 15, No. 1, p. 4183 (2024)
- [Chib 09] Chib, V. S., Rangel, A., Shimojo, S., and O'Doherty, J. P.: Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex, *J Neurosci*, Vol. 29, No. 39, pp. 12315–12320 (2009)
- [Foerde 15] Foerde, K., Steinglass, J. E., Shohamy, D., and Walsh, B. T.: Neural mechanisms supporting maladaptive food choices in anorexia nervosa, *Nature Neuroscience*, Vol. 18, No. 11, pp. 1571–1573 (2015)
- [Hare 09] Hare, T. A., Camerer, C. F., and Rangel, A.: Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System, *Science*, Vol. 324, No. 5927, pp. 646–648 (2009)
- [He 15] He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition (2015)
- [Iigaya 21] Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, K., and O'Doherty, J. P.: Aesthetic preference for art can be predicted from a mixture of low- and high-level visual features, *Nature Human Behaviour*, Vol. 5, No. 6, pp. 743–755 (2021)
- [Iigaya 23] Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, S., Cross, L., and O'Doherty, J. P.: Neural mechanisms underlying the hierarchical construction of perceived aesthetic value, *Nature Communications*, Vol. 14, No. 1, p. 127 (2023)
- [Liu 22] Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S.: A ConvNet for the 2020s (2022)
- [Motoki 20] Motoki, K. and Suzuki, S.: Extrinsic Factors Underlying Food Valuation in the Human Brain, *Front Behav Neurosci*, Vol. 14, p. 131 (2020)
- [Rangel 08] Rangel, A., Camerer, C., and Montague, P. R.: A framework for studying the neurobiology of value-based decision making, *Nature Reviews Neuroscience*, Vol. 9, No. 7, pp. 545–556 (2008)
- [Simonyan 15] Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition (2015)
- [Spinelli 21] Spinelli, S. and Monteleone, E.: Food Preferences and Obesity, *Endocrinol Metab (Seoul)*, Vol. 36, No. 2, pp. 209–219 (2021)
- [Suzuki 17] Suzuki, S., Cross, L., and O'Doherty, J. P.: Elucidating the underlying components of food valuation in the human orbitofrontal cortex, *Nat Neurosci*, Vol. 20, No. 12, pp. 1780–1786 (2017)
- [Yamins 14] Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., and DiCarlo, J. J.: Performance-optimized hierarchical models predict neural responses in higher visual cortex, *Proc Natl Acad Sci U S A*, Vol. 111, No. 23, pp. 8619–8624 (2014)