

NAIST-IS-MT1551070

修士論文

畳み込みニューラルネットワークを用いた画像修復の 品質評価に基づく画像修復結果の自動選出

南村 敏弥

2017年3月16日

奈良先端科学技術大学院大学
情報科学研究科

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

南村 敏弥

審査委員：

横矢 直和 教授	(主指導教員)
向川 康博 教授	(副指導教員)
佐藤 智和 准教授	(副指導教員)
中島 悠太 客員准教授	(副指導教員/大阪大学)
河合 紀彦 助教	(副指導教員)

畳み込みニューラルネットワークを用いた画像修復の 品質評価に基づく画像修復結果の自動選出*

南村 敏弥

内容梗概

現像した写真についての傷や、撮影時に意図せずに映った人物など、画像内に不要な領域が存在する場合がある。このような不要な領域を取り除き、取り除かれた領域である欠損領域を違和感なく埋める、画像修復と呼ばれる研究が盛んに行われている。従来手法として、画像内のテクスチャを事例として用いた手法が広く用いられている。これは、欠損領域と非欠損領域の間のパターン類似度に基づくエネルギー関数を作成し、それを最小化するような画素で欠損領域を埋めることで修復を行う手法である。しかし、この手法は初期値などのパラメータの設定によって修復結果が大きく変化し、一般に良い修復結果が得られるようなパラメータを事前に決定することができないという問題がある。

そこで本論文では、画像修復においてより良い修復結果を得ることを目的とし、畳み込みニューラルネットワークを用いて、複数の画像修復の結果から自動的に品質の良いものを選出する手法を提案する。提案手法では、画像修復結果の品質を評価するためのデータセットを用いて学習を行い、入力された修復画像の品質が良いか悪いかを出力する畳み込みニューラルネットワークを構成する。画像修復結果の品質を評価するために、本研究では2種のデータセットを構築し転移学習を行う。1つ目のデータセットは、元画像に対して無作為に欠損領域を指定し画像修復を適用したものであり、大量のデータの自動生成が可能である。提案手法では、このデータセットを用いて畳み込みニューラルネットワークを事前学習

*奈良先端科学技術大学院大学 情報科学研究科 修士論文, NAIST-IS-MT1551070, 2017年3月16日.

する。2つ目のデータセットは、手作業により欠損箇所を指定し、様々なパラメータによって画像修復を行った結果それぞれに対して、手作業で品質の高低をラベルを付与したものである。このデータセットを利用して畳み込みニューラルネットワークをファインチューニングすることにより、最終的な学習結果を得る。これにより、手作業での欠損領域の指定とラベルの付与によるデータが少ない場合でも、高い性能を達成することが期待できる。このネットワークを用いて修復結果を評価することにより、様々なパラメータにより修復された複数の画像の中から品質の良いものを自動的に選出することが可能である。実験では、学習済みのネットワークで画像修復結果の自動選出を行い、主観的に選出したものとどの程度一致するかを確認することで提案手法の有用性を示す。

キーワード

品質評価, 画像修復, 畳み込みニューラルネットワーク

Automatic Selection of Image Inpainting Results Based on Quality Evaluation Using Convolutional Neural Network*

Harumi Namura

Abstract

Undesirable objects may be present in images, such as scratches on printed images or unintentionally captured people. Image inpainting is a technique to remove such undesirable regions and fill the missing regions in a visually plausible way. For doing this, methods have been widely studied that copy some regions in the image as examples to the missing regions. These methods use an energy function based on the similarity between the missing regions and other regions, and fill the missing regions by minimizing this function. However, these methods have a problem that the quality of the inpainting results greatly varies depending on parameters of the methods, such as an initial value, and it is impossible to determine such parameter values that always give a good inpainting result.

This thesis proposes a method to evaluate inpainting results using a convolutional neural network (CNN) for automatically selecting a better image inpainting result from ones obtained with various parameters. To the best of our knowledge, a dataset that can be used to train the CNN has not been built; therefore in order to efficiently train our CNN, we build two datasets and perform transfer learning. The first dataset contains images with randomly set missing regions, together with their inpainting results. Since the missing regions are randomly

*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1551070, March 16, 2017.

set, we can automatically generate a large number of data. The second dataset contains images with manually specified missing regions. For this dataset, we apply an image inpainting method with various parameter values, and manually assign a label that represents either good or bad quality. We use this dataset to fine-tune our CNN so that its scores correlate with the labels. By doing this, we can train the CNN without using a vast amount of manually labeled data. By evaluating the inpainting results using this CNN, we can automatically select a result with good quality out of ones obtained with various parameters. In our experiments, we demonstrate the effectiveness of the proposed method by confirming how much our CNN is consistent with manually assigned labels.

Keywords:

quality evaluation, image inpainting, convolutional neural networks

目次

1. はじめに	1
2. 従来研究および本研究の位置づけ	4
2.1 画像修復に関する従来研究	4
2.1.1 輝度値の連続性を考慮した画像修復	4
2.1.2 画像内のテクスチャを事例として用いる画像修復	5
2.1.3 畳み込みニューラルネットワークを用いた画像修復	10
2.2 本研究の位置づけ	12
3. 畳み込みニューラルネットワークを利用した画像修復の品質評価に基づく 画像修復結果の自動選出	14
3.1 提案手法の概要	14
3.2 画像修復の手法	16
3.2.1 パターン類似度によるエネルギー関数の定義	16
3.2.2 エネルギー関数の最小化による画素値の更新	18
3.2.3 粗密法による局所解の回避と高速化	18
3.2.4 パラメータによる修復結果の変化	19
3.3 データセットの生成	22
3.3.1 事前学習用データセットの生成	22
3.3.2 ファインチューニング用データセットの生成	22
3.4 品質評価のための畳み込みニューラルネットワークの構成	24
3.5 畳み込みニューラルネットワークの学習	26
3.6 学習済みネットワークによる修復画像の選出	27
4. 実験と考察	28
4.1 畳み込みニューラルネットワークの学習実験	28
4.1.1 データセットの生成	28
4.1.2 畳み込みニューラルネットワークの学習	29
4.2 学習済みネットワークによる修復画像の選出実験	31

4.3 被験者による主観的評価実験	37
4.3.1 修復結果の手動による評価の検証	37
4.3.2 学習済みネットワークによる修復画像の選出の検証	40
5. まとめ	45
謝辞	46
参考文献	47

目 次

1	パラメータによる修復結果の違い	2
2	Bertalmio らの手法による画像修復結果	5
3	天野らの手法による画像修復結果	6
4	Criminisi らの手法による画像修復結果	7
5	Wexler らの手法による画像修復結果	8
6	Darabi らの手法による画像修復結果	9
7	Kawai らの手法による画像修復結果	10
8	Xie らの手法による画像修復結果	11
9	Pathak らの手法による画像修復結果	12
10	提案手法の概略	15
11	各領域の概略	17
12	元画像と与えた欠損	19
13	距離の重みに対する修復結果の変化	20
14	テクスチャの複雑さに対する修復結果の変化	20
15	パターン類似度の重みに対する修復結果の変化	21
16	PatchMatch の初期値の違いによる修復結果の変化	21
17	事前学習用データセットの例	23
18	手動で欠損領域を指定した例	24
19	20 通りのランダムパラメータで修復を行った例	25
20	畳み込みニューラルネットワークの構成	26
21	データセット生成に利用した画像の例	28
22	評価に使用したインターフェース	30
23	欠損を与えた画像の例	32
24	修復画像の評価例 1 (p : 修復画像の品質についてのスコア)	33
25	修復画像の評価例 2 (p : 修復画像の品質についてのスコア)	34
26	修復画像の評価例 3 (p : 修復画像の品質についてのスコア)	35
27	修復画像の評価例 4 (p : 修復画像の品質についてのスコア)	36
28	評価に使用したインターフェース	37

29	ラベルが高の画像を選択した人数が多い画像例 1	38
30	ラベルが高の画像を選択した人数が多い画像例 2	39
31	ラベルが高の画像を選択した人数が少ない画像例 1	39
32	ラベルが高の画像を選択した人数が少ない画像例 2	39
33	被験者全員が提案手法による修復画像を選択した例 1 (p : 修復画像の品質についてのスコア)	41
34	被験者全員が提案手法による修復画像を選択した例 2 (p : 修復画像の品質についてのスコア)	41
35	被験者全員が従来手法による修復画像を選択した例 1 (p : 修復画像の品質についてのスコア)	42
36	被験者全員が従来手法による修復画像を選択した例 2 (p : 修復画像の品質についてのスコア)	43
37	3人の被験者が提案手法による修復画像を選択した例 1 (p : 修復画像の品質についてのスコア)	43
38	3人の被験者が提案手法による修復画像を選択した例 2 (p : 修復画像の品質についてのスコア)	44

表 目 次

1	識別率	31
2	ラベルが高の画像を選択した人数とその枚数	38
3	提案手法による修復画像を選択した人数とその枚数	40

1. はじめに

近年、デジタルカメラやスキャナ、パーソナルコンピュータ等のデジタル機器の高性能化、低価格化が進み、デジタル画像が広く扱われるようになった。さらに、スマートフォンの普及によって、ブログやSNSを通じてインターネットにデジタル画像をアップロードすることが一般的に行われるようになった。ここで、アナログ画像をスキャンした際の経年劣化による傷、ほこりや、観光地で記念撮影した際に写り込んだ他の観光客など、見栄えやプライバシーの問題から画像中に取り除きたい領域が存在する場合がある。

このような不要な領域（以下、欠損領域）を違和感なく埋めることで、画像の価値を高める画像修復と呼ばれる研究が盛んに行われている。画像修復の既存の研究は、輝度値の連続性を考慮した手法 [1-5]、事例に基づいた手法 [6-29]、畳み込みニューラルネットワークを用いた手法 [30,31] に大別できる。輝度値の連続性を考慮した手法では、欠損領域の周辺の輝度値を、欠損領域の内側へ滑らかに伝播させていくことで修復を行う。この手法では細かいテクスチャを再現できないため、欠損領域が小さい場合は良い修復結果を得ることができるが、大きい場合は不鮮明な修復結果になるという問題がある。また、畳み込みニューラルネットワークを用いた手法では任意の形状の欠損領域を修復することが難しいという問題がある。

画像内のテクスチャを事例として用いる手法は、現在に至るまで広く研究されている。この手法では、欠損領域と、欠損領域以外の領域（以下、非欠損領域）の間のパターン類似度に基づき、非欠損領域のテクスチャを欠損領域にコピーすることで修復を行う。非欠損領域のテクスチャを用いるため、細かいテクスチャが再現でき、大きな欠損でも良好な修復結果を得やすい。しかし、パターン類似度に基づくエネルギー関数には、欠損領域の境界からの距離に関する項やテクスチャの複雑さに関する項などが含まれており、それぞれの項に対してパラメータを経験的に決定する必要がある。また、非欠損領域中の類似パターンの探索は時間のかかる処理であることから、ランダムな初期値によって画像中の探索位置を決定する PatchMatch [29] を利用した高速化が広く用いられている。これらのパラメータや初期値の設定によって、図1に示すように修復結果が大きく変化する



(a) 元画像



(b) 欠損を与えた画像



(c) 修復画像の例 1



(d) 修復画像の例 2

図 1: パラメータによる修復結果の違い

場合がある。しかし、これらの手法には、一般に良い修復結果が得られるようなパラメータを事前に決定することができないという問題がある。

本論文では、画像修復においてより良い修復結果を得ることを目的とし、畳み込みニューラルネットワークを用いて、事例に基づいた手法による画像修復の結果から、自動的に品質の良いものを選出する手法を提案する。提案手法では、画像修復結果の品質に関するデータセットを用いて学習を行い、入力された修復画像の品質が良いか悪いかを出力する畳み込みニューラルネットワークを構成する。画像修復結果の品質を評価するために、本研究では、品質について「高」、もし

くは「低」のいずれかのラベルを付与した画像からなるデータセットを構築し、これを用いて畳み込みニューラルネットワークを学習する。このネットワークで修復結果を評価することにより、様々なパラメータにより修復された複数の画像の中から品質の良いものを自動的に選出することが可能である。

以下、2章では、画像修復に関する関連研究と、本研究の位置付けについて述べる。3章では、本研究で利用する既存の手法の詳細と、畳み込みニューラルネットワークの構成、学習方法について述べる。4章では、畳み込みニューラルネットワークの学習実験の結果を示し、従来の結果の品質評価を行わない手法と、提案手法の主観的評価を行うことで、提案手法の有効性を示す。5章ではまとめと今後の展望について述べる。

2. 従来研究および本研究の位置づけ

本章では、画像修復に関する従来手法について概観し、本研究の位置付けについて述べる。

2.1 画像修復に関する従来研究

画像修復に関する従来研究は、輝度値の連続性を考慮した手法 [1–5]，事例に基づいた手法 [6–29]，畳み込みニューラルネットワークを用いた手法 [30, 31] に分類できる。以下では、各従来手法の特徴や問題点について述べる。

2.1.1 輝度値の連続性を考慮した画像修復

輝度値の連続性を考慮した手法は、欠損領域周辺から欠損領域内へ輝度値が滑らかに変化していれば違和感が少ない画像になるという考え方にに基づき、欠損領域周辺の画素値に基づいて欠損領域を補間することで修復を行う。

このような手法において、輝度値の連続性のみを考慮するとシーンの構造が再現されない。そこで、多くの手法ではエッジを保存するように修復を行う。前田ら [1] は、欠損領域の境界部でのエッジの有無を推定し、エッジが存在する場合には欠損領域内部へエッジを伸ばすような手法を提案した。この手法では欠損領域の境界から内部へ逐次的に修復を行うため、欠損が大きい場合にエッジが繋がらないという問題がある。また Masnou ら [2] は、欠損領域の周りの複数のエッジを自動的に予め対応させ、それらを直線的に結ぶことで修復を行う手法を提案した。この手法では、欠損領域の周りのエッジが複雑な場合、エッジを正しく対応付けることが難しい。またエッジを直線的に結ぶため、欠損領域周辺でエッジが非直線的なパターンを形成する場合、違和感のある修復結果となる。これらの問題に対して、偏微分方程式を用いた手法 [3] や、弾性方程式を用いた手法 [4]，オプティカルフローを応用する手法 [5] など、反復的にエッジを接続しながら修復を行うことで、違和感のない修復を目指す手法が提案されている。

輝度値の連続性を考慮した手法では、細かいテクスチャを生成することができないため、欠損領域が小さい場合には図2のように良好な修復結果を得ることが



(a) 欠損が与えられた画像



(b) 修復された画像

図 2: Bertalmio らの手法 [3] による画像修復結果

できるが、大きな欠損領域を修復した場合には不鮮明な修復結果となるという問題がある。

2.1.2 画像内のテクスチャを事例として用いる画像修復

画像内のテクスチャを事例として用いる手法は、欠損領域内に存在するべきテクスチャが非欠損領域に存在するテクスチャに類似するという仮定に基づき、非欠損領域のテクスチャを用いて修復を行う手法である。欠損領域が大きくても良好な結果を得やすいことから現在広く研究されている。この手法は特徴空間を用いる手法 [6–10]、逐次的に修復を行う手法 [11–20]、欠損領域全体で最適化を行う手法 [21–28] に分類することができる。

特徴空間を用いる手法では、欠損領域を含む一定範囲の領域や非欠損領域内の領域から特徴量を求め、特徴空間内で類似パターンを探索し、修復を行う。東海林ら [6] は非欠損領域内のフーリエ振幅スペクトルを特徴量として用いる手法を提案した。この手法はフーリエ振幅スペクトルの位置不変性を仮定するので、周期的なテクスチャパターンを持つ画像では良好な修復結果を得ることができるが、そうではない場合は難しい。また Hirani ら [7] は、空間領域と周波数領域の両方を用いて修復を行う手法を提案した。しかしこの手法では欠損領域の修復に用いる領域を手動で指定する必要があり、非欠損領域で多様なテクスチャを持つ場合には修復が難しい。



(a) 欠損が与えられた画像



(b) 修復された画像

図 3: 天野らの手法 [8] による画像修復結果

また，画像から得られる固有ベクトルを用いた手法が提案されている．天野らは，画像の局所領域が，同じ画像の別の局所領域と類似するという性質を表す自己相関性の高い画像を対象とした手法である BPLP 法 [8]，及びその改良である kBPLP 法 [9] を提案した．この手法では，画像の非欠損領域から学習サンプルとして複数の局所領域を切り出し，KL 展開などによって固有ベクトル群を生成し，これらを一次結合することで修復を行う．図 3 に示すように，自己相関性の高い画像に対して良好な修復結果を得ることができる．また，井添ら [10] は，画像の局所性とフラクタル性に着目し，画像から生成したベクトルにカーネル主成分分析を行う事で修復を行う Kernel FID 法を提案した．この手法も非欠損領域のテクスチャを利用するため，細かいテクスチャを再現することができる．

上記のような特徴空間を利用する手法は，欠損領域周辺の特徴量を用いるため，欠損領域にウィンドウを当てはめた際にその中に必ず非欠損領域を含まなければならないという制約がある．よって，1つのウィンドウに収まらないような大きな欠損領域をもつ画像に対しては適用が難しい．

この問題に対して，逐次的に修復を行う手法が提案されている．この手法では，欠損領域の境界上の局所領域のテクスチャパターンについて，非欠損領域内から類似する部分を探索し，最も類似するものを逐次的に合成することで修復を行う．図 4 の様に大きな欠損領域の場合でも良好な修復結果を得ることができる．Efros ら [11] は，テクスチャ合成の分野で用いられてきた手法を画像修復に応用する手



(a) 人物が写っている画像



(b) 人物が除去された画像

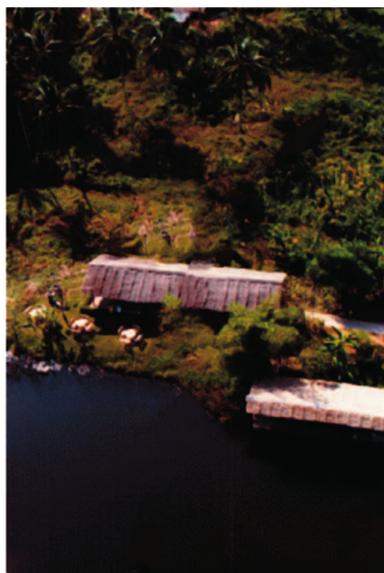
図 4: Criminisi らの手法 [16] による画像修復結果

法を提案した。また，Bertalmio ら [12] はこの手法を発展させ，輝度値の連続性を考慮した手法 [3] と組み合わせた手法を提案した。これらの手法は，欠損領域の境界から欠損領域の内部へと逐次的に合成を行い，一度合成された箇所は固定するという方針で修復を行うため，修復にかかる時間が短い，修復の品質が合成の順序に大きく依存するという問題がある。この問題に対して，修復結果の違和感が少なくなるような合成順序を得るため，以下のような様々な基準を用いた手法が提案されている。

- テクスチャの類似度 [13]
- invers matte [14]
- 決定済みの画素の数 [15]
- 決定済みの画素数とエッジの強さ [16–18]
- 決定済みの画素数とエッジの強さと類似度 [19]
- テクスチャの希少度 [20]



(a) 人物が写っている画像



(b) 人物が除去された画像

図 5: Wexler らの手法 [24] による画像修復結果

しかし、これらの手法においても、非欠損領域で複雑なテクスチャ構造を持つ画像では、修復結果にテクスチャの不連続が生じやすいという問題がある。

逐次的に合成を行う手法の問題点を解決するために、欠損領域全体でテクスチャを最適化することによって修復を行う手法が提案されている。Komodakis ら [21] は、Belief Propagation を応用した Priority-BP を用いて欠損領域の最適な修復順を得ることで、不連続なテクスチャの発生を抑える手法を提案した。また Pritch ら [22] は、欠損領域内の画素と非欠損領域内の画素を対応付ける Shift Map をグラフカットを用いて最適化する事で修復を行う手法を提案した。Allène ら [23] は、欠損領域を様々な形に分割し、最適となるテクスチャの組み合わせを探索し合成を行う手法を提案した。Wexler ら [24] は、欠損領域と非欠損領域のパターン類似度に基づいた、画像の尤もらしさを表すエネルギー関数を欠損領域全体で定義し、これを欠損領域全体に関して最小化するように画素を決定することで修復を行う手法を提案した。このような画像全体での最適化を行う手法では、図 5 に示すように、大きい欠損領域であっても良好な修復結果を得ることができる。

しかし、これらの手法では、パターン類似度の計算にパターン間の輝度値の差



(a) 欠損が与えられた画像



(b) 修復された画像

図 6: Darabi らの手法 [25] による画像修復結果

の二乗和である SSD などの指標を用いるため、一見類似するようなパターンが非欠損領域内に存在しても、明度や見え方に変化がある場合にそのパターンを利用できず、良好な修復結果が得られないという問題がある。この問題を解決しより多様な画像で良好な修復結果を得るために、パターンの明度変化や幾何学的変形を考慮することで利用できるパターンを増加させる手法が提案されている。Darabi ら [25] は、パターン類似度に基づくエネルギー関数を拡張し、パターンの回転・拡大縮小・反転を考慮した手法を提案した。これにより図 6 のような回転やスケール変化のある物体上の欠損領域でも、パターンの変形によって良好な修復結果を得ることができる。Kawai らは、パターン類似度に基づくエネルギー関数 [24] を拡張し、明度変化に対応した手法 [26] およびパターンの対称性を考慮した手法 [27] を提案した。これにより図 7 のように、画像内に類似パターンが存在しない場合でも、対称性を考慮することでパターンを生成し、良好な修復結果を得ることができる可能性がある。また Huang らが提案した手法 [28] では、画像内の平面を検出しその向きに応じた透視投影変換をその平面領域内のパターンに施すことで、透視投影歪みが生じパターンの見え方が異なる場合でも良好な修復結果を得ることができる。

また、画像内のテクスチャを事例として用いる手法では、類似パターンの探索を非欠損領域全体で行うと計算量が非常に大きくなり、処理に時間がかかるとい



(a) 欠損が与えられた画像



(b) 修復された画像

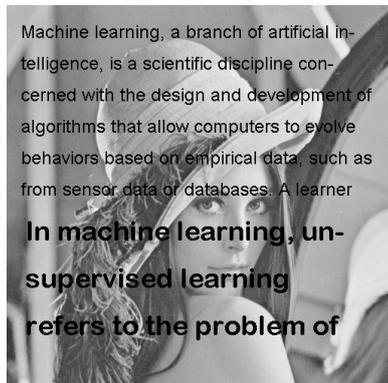
図 7: Kawai らの手法 [27] による画像修復結果

う問題がある．この問題に対し，Barnes ら [29] は類似パターンの探索の高速化手法として PatchMatch を提案した．PatchMatch では，まず欠損領域内の各パターンに対して，非欠損領域内のランダムな位置のウィンドウを類似パターンの初期値として与える．次に，ある 1つのパターンについて，隣接するパターンに与えられている類似パターンの近傍から類似パターンを探索する．このような手順で探索範囲を限定することで計算コストを大きく削減しており，Kawai らの手法 [27] や Huang らの手法 [28] などに利用されている．

このように，欠損領域全体で最適化を行う手法は様々な画像に対して良好な修復結果を得ることができる．しかしこれらの手法は，エネルギー関数内にパターン間の距離やテクスチャの複雑さなどの項があり，それぞれの重みがパラメータとして存在する．これらの値や，PatchMatch の初期値によって，修復結果が大きく変化するという問題がある．異なる画像を同一のパラメータによって修復した際に，良好な修復結果を得ることができる場合とそうでない場合があり，一般に良い修復結果が得られるようなパラメータを事前に決定することは難しい．

2.1.3 畳み込みニューラルネットワークを用いた画像修復

近年のディープラーニングの普及と共に，畳み込みニューラルネットワークを用いた手法が提案されている．Xie ら [30] は，元画像と，それにノイズや文字を



(a) 欠損が与えられた画像



(b) 修復された画像

図 8: Xie らの手法 [30] による画像修復結果

重畳した画像からなるデータセットを作成し、元画像を教師画像として学習したネットワークを用いることで修復を行う手法を提案した。この手法では、文字などが重畳された未知の画像を学習済みのネットワークに入力すると、欠損領域の指定なしで図 8 に示すような結果を得ることができる。しかしこの手法は、文字などの小さい欠損領域の修復を対象としているため、大きな欠損領域に対して良好な結果を得ることは難しい。

また Pathak ら [32] は、画像生成などにおいて近年盛んに利用されている GAN (Generative Adversarial Network [31]) と呼ばれるネットワーク構造を応用した手法を提案している。GAN は、画像を生成する Generator と、入力画像が Generator によって生成されたものか実際に撮影された画像かを判定する Discriminator の 2 つのネットワークからなる。2 つのネットワークを同時に学習することで、Discriminator の出力に基づいて Generator のパラメータを更新できるため、生成される画像が実際に撮影された画像に近づくという手法である。Pathak らの手法 [32] では、GAN の Generator において、欠損の周辺の画像を入力として与えて画像生成を行うことで、図 9 のような、周辺の画像と一貫性のある修復結果を得ることができる。しかしこの手法では、画像全体のサイズと欠損の形状や位置、大きさが固定されており、任意の欠損に対応出来ないという問題がある。



(a) 欠損が与えられた画像



(b) 修復された画像

図 9: Pathak らの手法 [32] による画像修復結果

2.2 本研究の位置づけ

2.1 節で概観したように、画像修復に関する様々な手法がこれまでに提案されている。輝度値の連続性を考慮した手法では、欠損領域が小さい画像を対象とする場合には良好な修復結果が得られるが、テクスチャが再現できないため、大きな欠損領域に対してはボケが発生し良好な修復結果が得られないという問題がある。一方、画像内のテクスチャを事例として用いた手法では、より大きな欠損領域を持つ画像でも良好な修復結果を得る事ができる。また、パターンの拡張を考慮することにより、より多様な画像に対して良好な結果となる可能性を高めることができる。しかし、エネルギー関数内の各項の重みや、PatchMatch の初期値によって図 1 のように修復結果が大きく変化するという問題がある。また、畳み込みニューラルネットワークを用いた手法では、任意の形状の欠損を修復するのは難しい。

本論文では、任意の形状で修復を行うことができる画像内のテクスチャを事例として用いる手法に基づき、畳み込みニューラルネットワークを用いて、複数の画像修復の結果から自動的に品質の良いものを選出する手法を提案する。画像修復を行った際に品質の悪い修復結果が得られた場合、不自然なエッジの途切れやテクスチャの乱れが発生していることが多い。GAN による画像生成 [31] や画像修復 [32] では、畳み込みニューラルネットワークを用いて、入力画像が生成された

ものか実際に撮影されたものか判定を行う。提案手法ではこれを参考にして、畳み込みニューラルネットワークを用いることで、画像修復結果が自然かどうかを判定することができると考え、修復画像を入力しその品質を出力するネットワークを構築する。異なるパラメータを用いて修復を行った複数の修復画像を入力し、評価値を比較することによって、良好な修復結果なものを選出することができる。これによって、従来手法で画像修復を行う際に、固定されたパラメータによる修復では良好な修復結果が得られない画像でも、良好な修復結果が得られることが期待できる。

3. 畳み込みニューラルネットワークを利用した画像修復の品質評価に基づく画像修復結果の自動選出

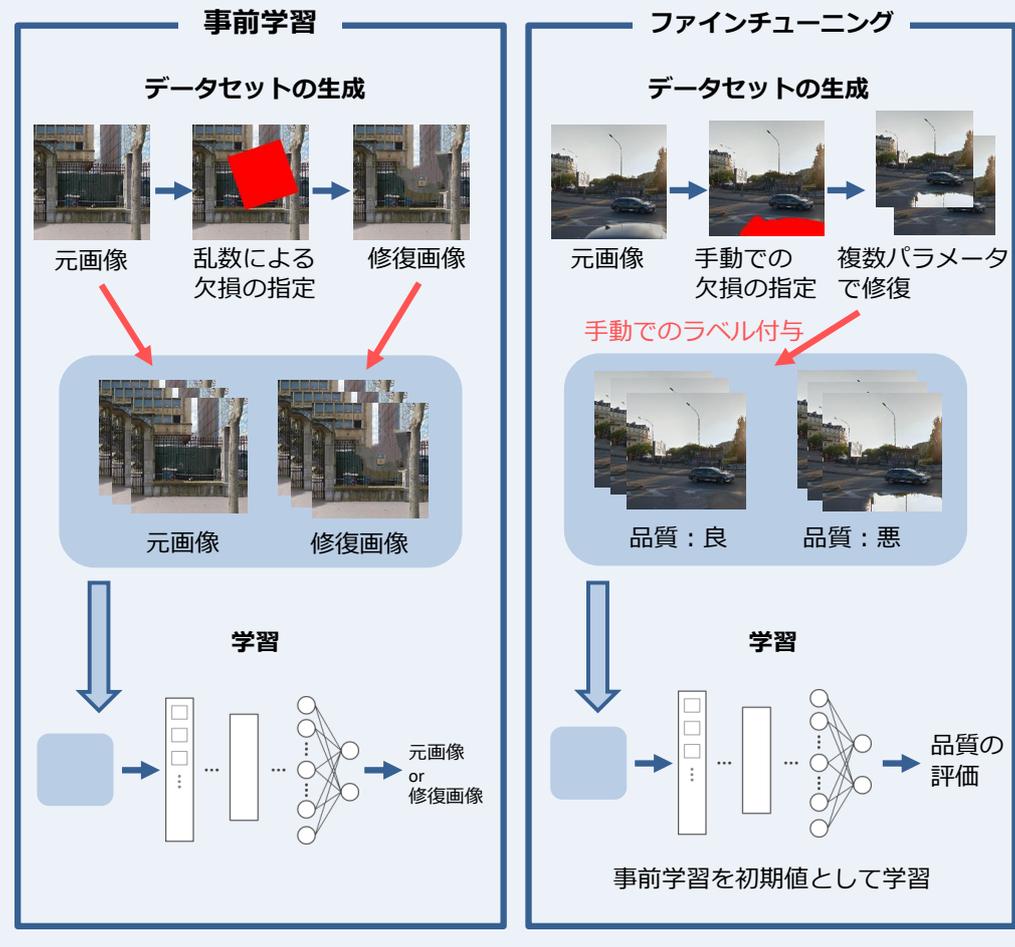
3.1 提案手法の概要

図 10 に提案手法の概略を示す。提案手法は、画像修復の品質評価のための畳み込みニューラルネットワークの学習、学習済みネットワークによる修復画像の選出からなる。

畳み込みニューラルネットワークの学習では、修復画像の品質を評価するネットワークの学習に必要なデータセットの生成を行う。データセットの作成では、修復の品質について「高」、もしくは「低」のいずれかのラベルを付与した修復済み画像からなるデータセットを構築する。この際、手動でのラベルの付与は時間がかかるため、大量のデータを用意するのは困難である。そこで、提案手法では転移学習を行う。転移学習は、あるタスクのためのネットワークの学習に、他のタスクのために学習されたネットワークを利用する手法である。事前学習として本来のタスクと近く、大量のデータを用意できる別のタスクのデータセットで学習を行い、そのネットワークを初期値としてファインチューニングを行うことで、データ数が少ない場合でも良い結果が得やすいことが知られている [33]。これに基づき、提案手法では修復画像と元画像の識別をタスクとして事前学習を行い、修復画像の品質の評価をタスクとしてファインチューニングを行う。このため、2種類のデータセットを構築する。1つ目は、乱数による欠損領域の指定によって自動生成した、事前学習用のデータセットである。これは、欠損領域を乱数によってランダムに与え修復を行った画像および欠損を与える前の画像からなる。2つ目は、欠損領域の手動指定によって手動生成した、ファインチューニング用のデータセットである。これは、手動で欠損領域を指定し、それぞれの画像に対して複数のパラメータで修復を行い、修復結果の中から手動で品質のラベルを付与したものである。その後、それぞれのデータセットを用いて事前学習及びファインチューニングを行い、入力した修復画像の品質を評価するネットワークを構築する。

学習済みネットワークによる修復画像の選出では、修復を行いたい画像を複数

畳み込みニューラルネットワークの学習



学習済みネットワークによる修復画像の選出

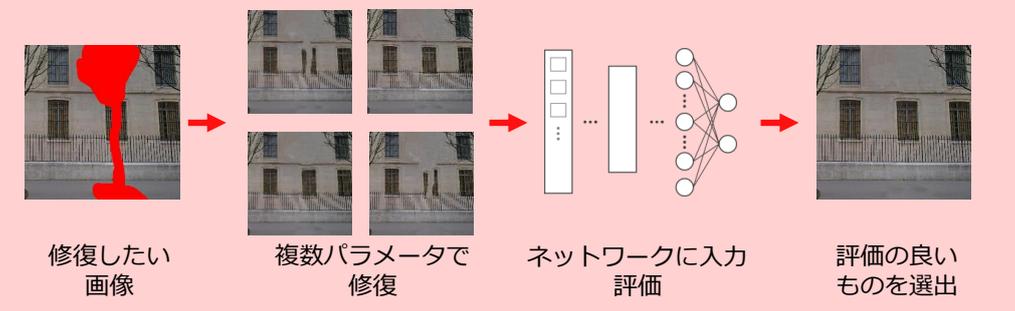


図 10: 提案手法の概略

組のパラメータで修復し，学習済みネットワークによってそれぞれの画像を評価し，評価の良いものを選出する．

以下，3.2節で提案手法で用いる画像修復の手法の詳細について，3.3節でデータセットの生成について，3.4節で畳み込みニューラルネットの構成について，3.5節でネットワークの学習について，3.6節で修復画像の選出について述べる．

3.2 画像修復の手法

この節では，本研究で画像修復手法として利用する，欠損領域全体で最適化を行う手法の一つである Kawai らの手法 [27] に基づいた手法の詳細について述べる．この手法では，欠損領域と非欠損領域との間のパターン類似度に基づくエネルギー関数を定義し，欠損領域内の画素値を，エネルギーを最小化するように更新することで修復を行う．また，良い修復結果が得やすいように，パターンに対して明度変化・局所性・対称性を考慮している．以下，詳細について述べる．

3.2.1 パターン類似度によるエネルギー関数の定義

欠損領域を含む画像において，図 11 のように欠損領域を Ω ，画像内のある画素を中心とした一定サイズのウインドウ W を考えた際に，その W 中に Ω が含まれるような中心画素の集合である領域を Ω' ， Ω' 以外の非欠損領域を Φ とする．欠損領域と非欠損領域との間のパターン類似度に基づくエネルギー関数は，領域 Ω' 内の画素 \mathbf{x}_i と非欠損領域 Φ 内の画素 \mathbf{x}_j 周辺のパターン類似度に基づき以下のよう定義される．

$$E = \sum_{\mathbf{x}_i \in \Omega'} w_{\mathbf{x}_i} \{SSD(\mathbf{x}_i, \mathbf{x}_j, \mathbf{T}_{\mathbf{x}_i \mathbf{x}_j}) + \lambda SD(\mathbf{x}_i, \mathbf{x}_j)\} \quad (1)$$

ここで $SSD(\mathbf{x}_i, \mathbf{x}_j, \mathbf{T}_{\mathbf{x}_i \mathbf{x}_j})$ は領域 Ω' 内の画素を中心とするパターンと，非欠損領域 Φ 内の画素を中心とするパターン間の類似度に関する項， $SD(\mathbf{x}_i, \mathbf{x}_j)$ はテクスチャの局所性に関する項を表す．また， λ は SSD 項と SD 項の相対的な重みである． $w_{\mathbf{x}_i}$ は領域 Ω' 内の各画素に与えられた重みであり，エネルギー最小化のため

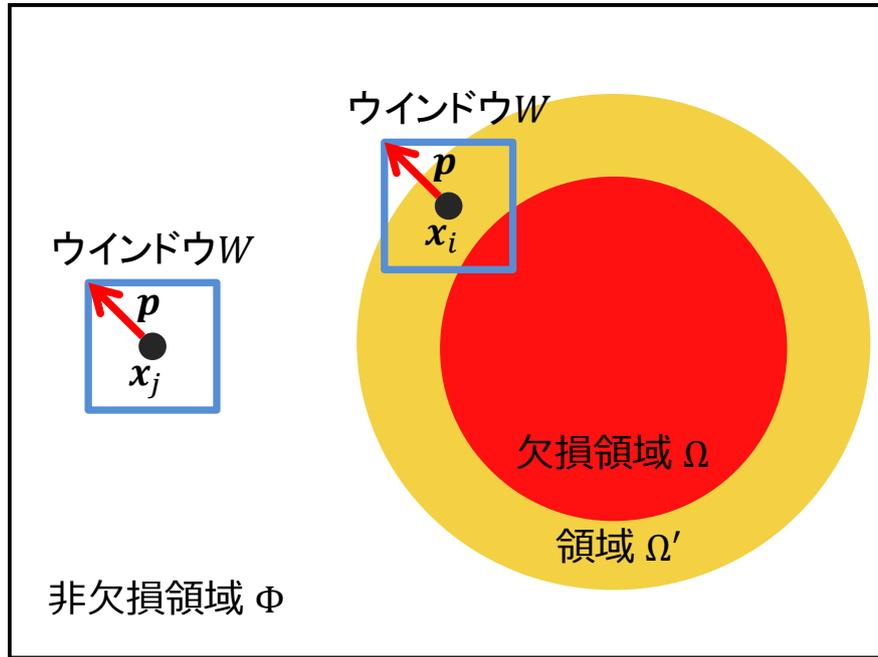


図 11: 各領域の概略

めの反復処理が行われる度に以下の式を用いて更新される.

$$w_{\mathbf{x}_i} = C(\mathbf{x}_i)S(\mathbf{x}_i)^\nu K(\mathbf{x}_i)^\mu \quad (2)$$

ここで, $C(\mathbf{x}_i)$ は

$$C(\mathbf{x}_i) = C^{-d_{\mathbf{x}_i}} \quad (3)$$

で表される, 定数 C と欠損領域 Ω から x_i までの最短距離 $d_{\mathbf{x}_i}$ に基づき計算されるウインドウの信頼度を, $S(\mathbf{x}_i)$ は領域 Ω' 内の画素 \mathbf{x}_i を中心とするウインドウとそれに対応するウインドウ間のパターン類似度, $K(\mathbf{x}_i)$ は Sobel フィルタにより計算される \mathbf{x}_i 周辺のテクスチャの複雑さをそれぞれ表す. また, ν, μ は各項の相対的な重みを表す.

式 (1) の $SSD(\mathbf{x}_i, \mathbf{x}_j, \mathbf{T}_{\mathbf{x}_i\mathbf{x}_j})$ は以下のように定義される.

$$SSD(\mathbf{x}_i, \mathbf{x}_j, \mathbf{T}_{\mathbf{x}_i\mathbf{x}_j}) = \sum_{\mathbf{p} \in W} \{I(\mathbf{x}_i + \mathbf{p}) - \alpha_{\mathbf{x}_i\mathbf{x}_j} I(\mathbf{x}_j + \mathbf{T}_{\mathbf{x}_i\mathbf{x}_j}\mathbf{p})\}^2 \quad (4)$$

ここで, $I(\mathbf{x}_i), I(\mathbf{x}_j)$ はそれぞれ画素 $\mathbf{x}_i, \mathbf{x}_j$ での画素値を表す. \mathbf{p} はウインドウ W 内での変位ベクトル, $\alpha_{\mathbf{x}_i\mathbf{x}_j}$ はテクスチャパターンの明度補正係数, $\mathbf{T}_{\mathbf{x}_i\mathbf{x}_j}$ はテ

クスタチャパターンの幾何学変換のための座標変換行列を表す。ここで、明度補正係数 $\alpha_{\mathbf{x}_i \mathbf{x}_j}$ は、 \mathbf{x}_i 周辺の画素と \mathbf{x}_j 周辺の画素の平均画素値の比から以下のように定義される。

$$\alpha_{\mathbf{x}_i \mathbf{x}_j} = \frac{\sqrt{\sum_{\mathbf{p} \in W} I(\mathbf{x}_i + \mathbf{p})^2}}{\sqrt{\sum_{\mathbf{p} \in W} I(\mathbf{x}_j + \mathbf{p})^2}} \quad (5)$$

明度補正係数によって、領域 Φ 内のパターンの明度を領域 Ω' の明度に合わせて補正するが、大きな輝度値の変化を線形的に近似すると違和感のある画像が生成されやすいため、値を一定の範囲に限定する。また、 $\mathbf{T}_{\mathbf{x}_i \mathbf{x}_j}$ によってパターンを変形することで対称性を考慮する。

また、式 (1) の $SD(\mathbf{x}_i, \mathbf{x}_j)$ は、シグモイド関数を用いて以下のように定義される。

$$SD(\mathbf{x}_i, \mathbf{x}_j) = \frac{N_W}{1 + e^{\{-K_0(\|\mathbf{x}_i - \mathbf{x}_j\|) - X_0\}}} \quad (6)$$

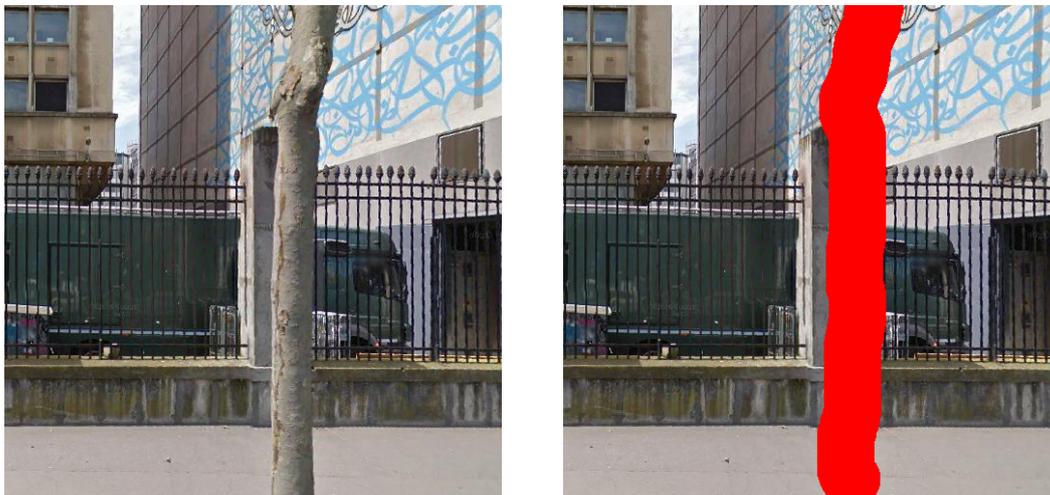
ここで、 K_0 、 X_0 は定数、 N_W はウインドウ内の画素数を表す。

3.2.2 エネルギー関数の最小化による画素値の更新

式 (1) で定義したエネルギー E が収束するまで、類似パターンの探索と画素値の更新を繰り返し、欠損領域内の各画素の値および対応する非欠損領域内の画素位置を決定する。

3.2.3 粗密法による局所解の回避と高速化

エネルギー関数の最小化は Greedy Algorithm の枠組みを用いているため、局所解問題が存在する。この手法では、効率的に大域最適解に近い値を得るため、多重スケール処理を行う。この処理では、まず一定のスケールになるまで画像サイズを半分ずつ縮小した画像ピラミッドを生成する。そして最小のスケールから順に、エネルギー関数の最小化を行い、得られた結果を1つ大きなスケールでの初期値として投影する。この処理を元のサイズになるまで繰り返し行うことで、大域最適解に近い解を得ることができ、各スケールでエネルギーの低い初期値を



(a) 元画像

(b) 欠損を与えた画像

図 12: 元画像と与えた欠損

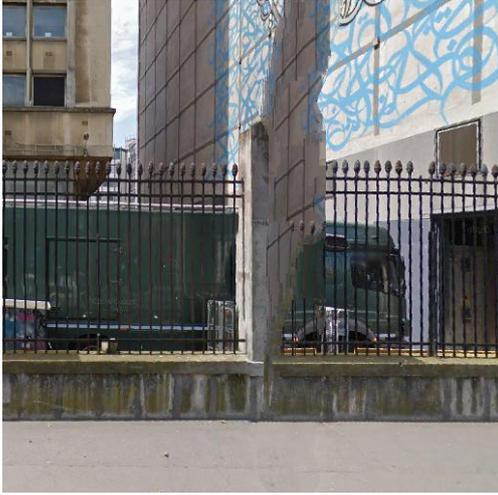
利用することができるので処理の高速化にも繋がる。また，PatchMatch [29] を利用して計算コストを削減する。

3.2.4 パラメータによる修復結果の変化

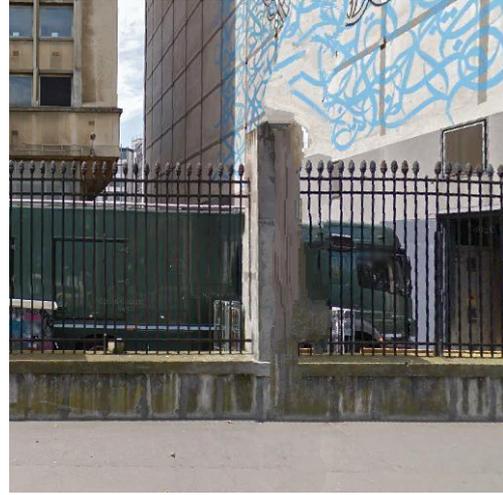
この手法では，式 (2) の ν や μ ，式 (3) の C といったパラメータの値や，PatchMatch の初期値によって，得られる修復結果が大きく変化する。すべての画像に対して良い修復結果が得られる様なパラメータは判明しておらず，経験的に以下の値が使われている。

- 距離： $C = 1.3$
- テクスチャの複雑さ： $\mu = 0.9$
- パターン類似度： $\nu = 0.6$

それぞれの値によって修復結果が変化する様子を，図 12 に示す画像に対して修復を行った場合を例として，それぞれ1つのパラメータを変化させ，他のパラメータを上記の値に固定し修復した結果を図 13，図 14，図 15，図 16 に示す。

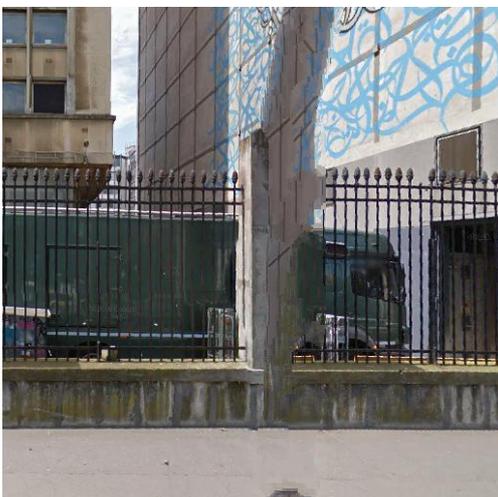


(a) $C = 1$

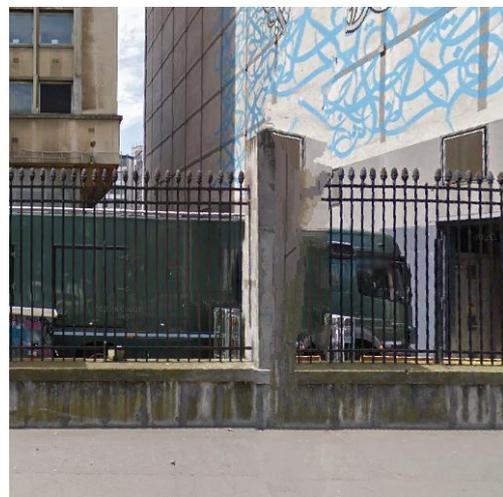


(b) $C = 2$

図 13: 距離の重みに対する修復結果の変化

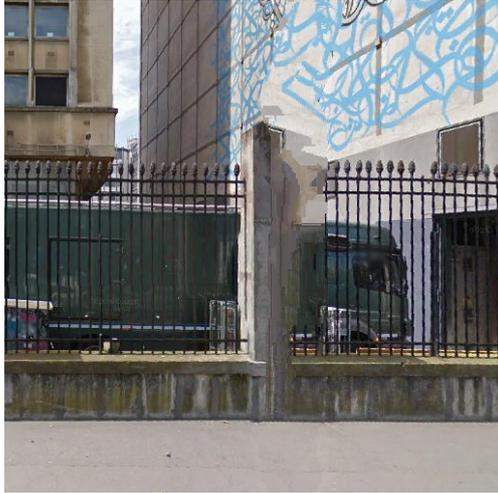


(a) $\mu = 0$



(b) $\mu = 1.5$

図 14: テクスチャの複雑さに対する修復結果の変化



(a) $\nu = 0$



(b) $\nu = 1.5$

図 15: パターン類似度の重みに対する修復結果の変化

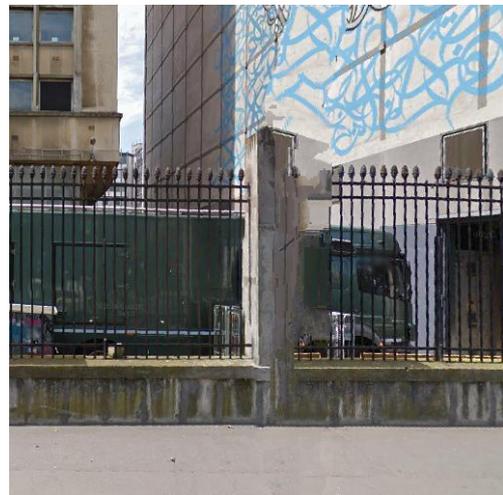
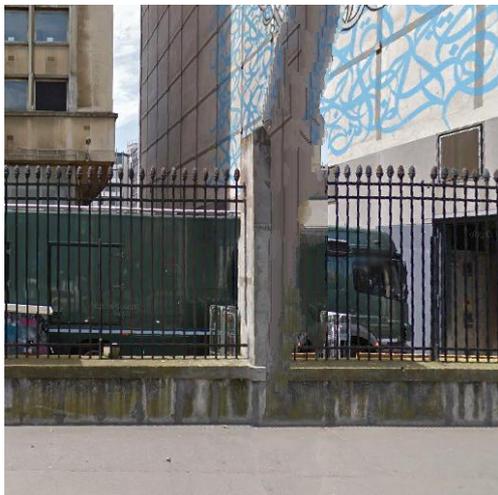


図 16: PatchMatch の初期値の違いによる修復結果の変化

3.3 データセットの生成

提案手法での畳み込みニューラルネットワークの学習では、2種類のデータセットを生成する。1つ目は修復画像と元画像の識別をタスクとした事前学習用のものである。これは、欠損の無い画像と、乱数により欠損領域を指定し修復を行った画像からなる。2つ目は修復画像の品質の評価をタスクとしたファインチューニング用のものである。これは、品質について「高」、もしくは「低」のいずれかのラベルを付与した修復画像からなる。

以下、それぞれの生成方法の詳細について述べる。

3.3.1 事前学習用データセットの生成

事前学習用データセットは、欠損を与えていない元画像と、乱数によって欠損領域を指定し修復を行った画像からなる。このデータセットを用いて、事前学習として修復を行った画像と元画像の分類を行う。元画像を修復結果の品質が高いもの、修復を行った画像を修復結果の品質が低いものとする、品質評価に関連があると考えられる。

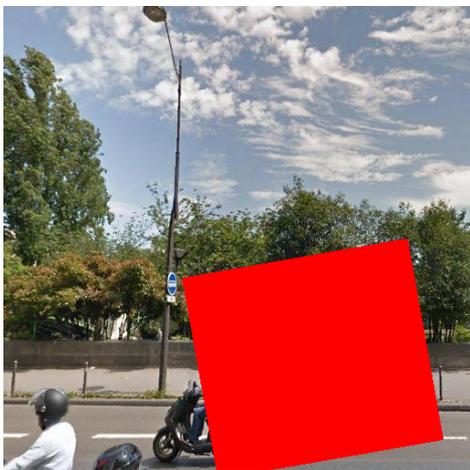
この生成では、まず用意した画像に、乱数によって欠損領域を指定する。欠損は矩形、正方形、正円、楕円の4種類のうち1つを、画像内に収まるような大きさ、位置になるような範囲で設定した乱数によって指定する。次に、この画像を、3.2節で述べた手法を用いて修復する。乱数によって欠損領域を指定し、修復を行った例を図17に示す。これらの欠損を指定する前の元画像と、修復を行った後の画像それぞれに正解のラベルを付与してデータセットを生成する。

3.3.2 ファインチューニング用データセットの生成

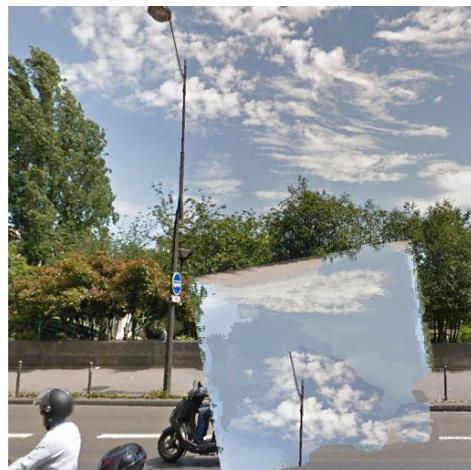
ファインチューニング用データセットは、手動で欠損領域を指定し、複数のパラメータによって修復した画像の中から品質の良いものと悪いものを手動で抽出しラベル付けした画像からなる。このデータセットを用いて、修復された画像の品質評価を行う。



(a) 元画像



(b) 欠損を与えた画像



(c) 修復結果

図 17: 事前学習用データセットの例



(a) 元画像

(b) 欠損を与えた画像

図 18: 手動で欠損領域を指定した例

この生成では、まず用意した画像に手動で欠損領域を指定する。指定した例を図 18 に示す。手動で欠損領域を指定した画像に対して、3.2 節で述べた手法を用いて、複数のランダムなパラメータによって修復を行い、複数通りの修復結果を得る。本手法では 20 組のランダムなパラメータを生成し、それぞれの組で同じ画像に対して画像修復を行った。その例を図 19 に示す。つぎに、得られた 20 通りの修復画像を確認する。この中に品質が良いと思えるものがあれば 1 枚選択し、品質について高のラベルを付与してデータセットに加える。また、悪いと思えるものがあれば 1 枚選択し、品質について低のラベルを付与してデータセットに加える。

3.4 品質評価のための畳み込みニューラルネットワークの構成

図 20 に、提案手法で用いる畳み込みニューラルネットワークの構成を示す。ネットワークは畳み込み層と全結合層からなる。畳み込み層は全 6 層であり、各層の活性化関数を Leaky ReLU 関数とする。また、全結合層は 6 層目の畳み込み層の出力を入力とし、活性化関数を Softmax 関数とする。2 つの出力は、事前学習ではそれぞれ入力画像が元画像であるか修復画像であるかの確率を、フィインチャー

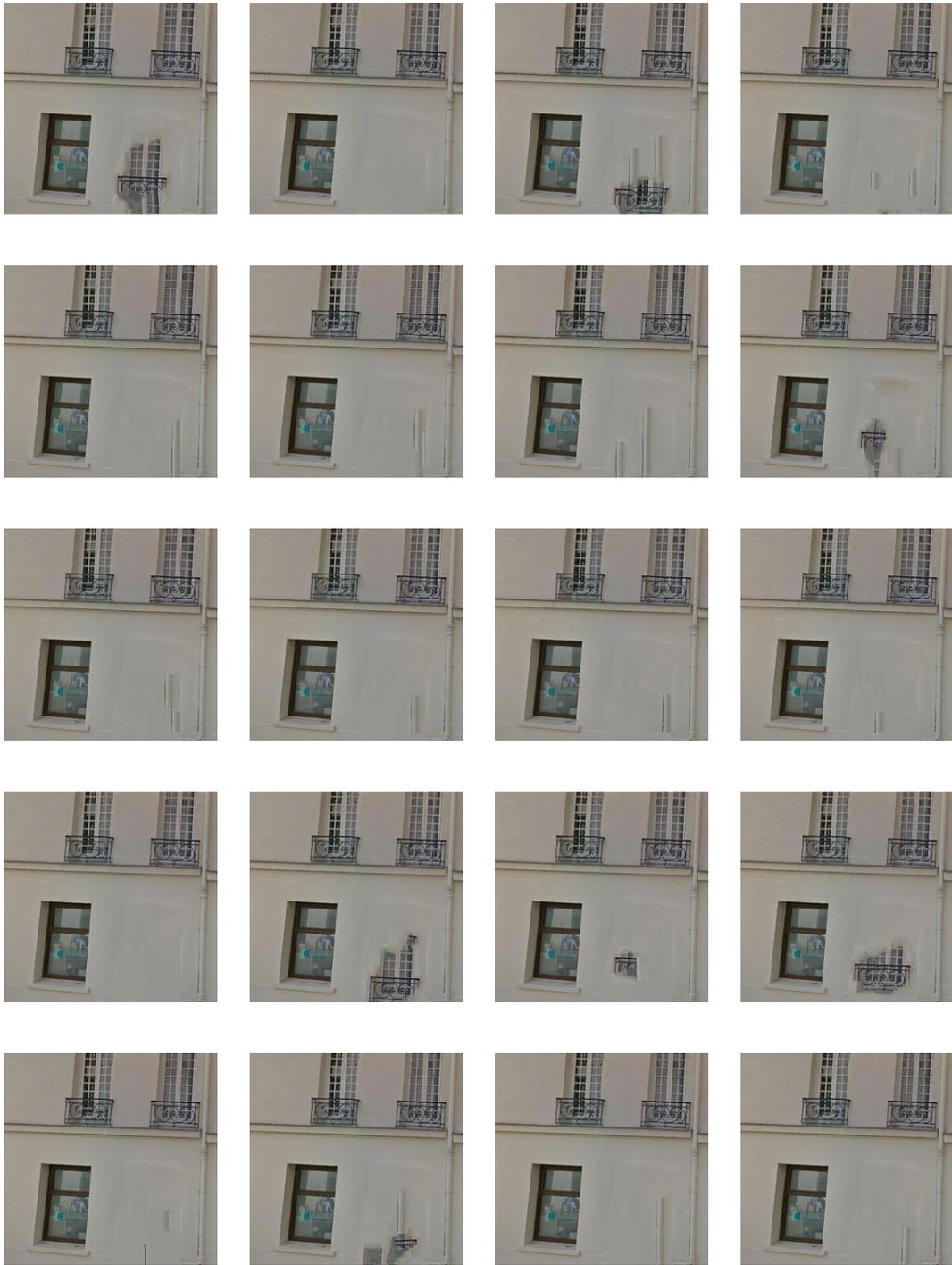


図 19: 20 通りのランダムパラメータで修復を行った例

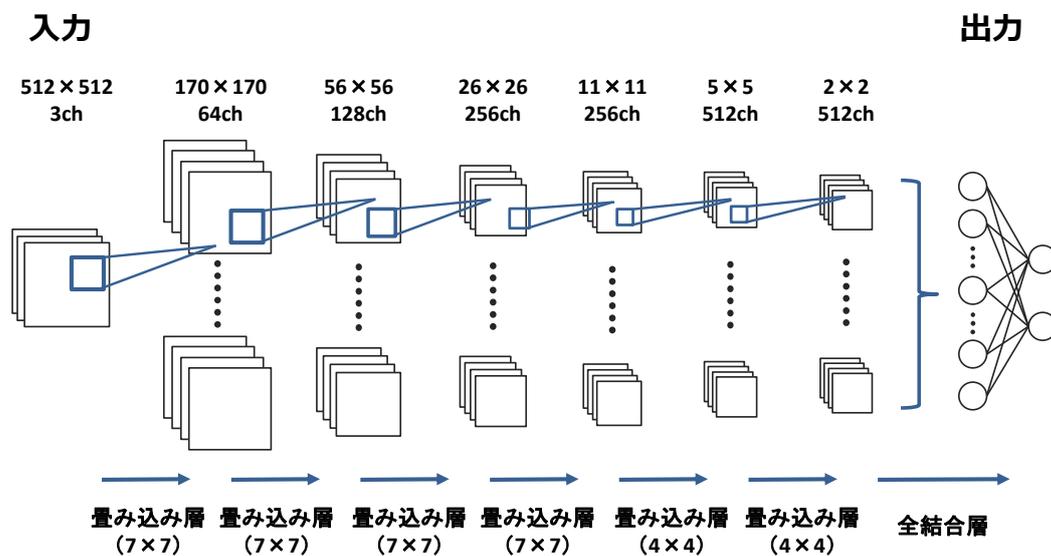


図 20: 畳み込みニューラルネットワークの構成

ニングでは修復画像の品質の高低の確率となる。

3.5 畳み込みニューラルネットワークの学習

畳み込みニューラルネットワークの学習では、まず 3.3.1 節で述べた事前学習用データセットによって、修復画像と元画像の分類のための学習を行う。その後、事前学習で得たネットワークを初期値として、3.3.2 節で述べたファインチューニング用データセットによって修復画像の品質評価のための学習を行う。この際、各畳み込み層で Batch Normalization [34] を行う。また、過学習を防ぐため全結合層に Dropout [35] を適用する。学習時には各データセットの画像を入力、付与されたラベルを教師とし、Cross Entropy によって誤差関数を定義し、確率的勾配法によってネットワークのパラメータを求める。

3.6 学習済みネットワークによる修復画像の選出

畳み込みニューラルネットワークの学習後，この学習済みネットワークを用いて修復画像の選出を行う．まず，修復対象となる，欠損領域を与えられた画像を複数組のパラメータにより修復する．これらを学習済みネットワークに入力し，品質の高低それぞれの確率を出力する．品質が高い確率をスコアとし，このスコアが一番高いものを最も品質の良い画像として選出する．

4. 実験と考察

提案手法の有効性を示すために，実験ではまず，3章で述べた手法により2種類のデータセットを作成し，生成したデータセットを用いて畳み込みニューラルネットワークを学習した上で，学習済みネットワークによる修復画像の評価を確認した．次に被験者による主観的評価を行うことで，提案手法の有用性を検証した．以下，実験の概要と評価結果，考察について述べる．

4.1 畳み込みニューラルネットワークの学習実験

4.1.1 データセットの生成

今回の実験では，提案手法のデータセットを生成するための元画像として Dorschら [36] が使用した，Google Street View から取得したパリの町並みのデータセットを使用した．このデータセットは， 936×537 ピクセルの画像 12,782 枚から構成される．図 21 に例を示す．



図 21: データセット生成に利用した画像の例

まず、3.3.1節で述べた事前学習用データセットの生成を行った。まず元のデータセットの画像1枚につき、ランダムに512×512の画像を8枚切り出した。このそれぞれに乱数によって欠損を指定し、3.2節で述べた手法を用いて修復を行った。これによって、元画像と修復画像のペア102,256組、合計204,512枚の画像を含む事前学習用データセットを生成した。

次に、3.3.2節で述べたファインチューニング用データセットの生成を行った。まず元のデータセットの画像1枚につき、512×512の画像を2枚切り出した。このそれぞれに手動で欠損領域を指定し、3.2節で述べた手法を用いて、20組のランダムなパラメータで修復を行った。なおこの際各パラメータは以下のような設定とした。

- 距離： $C \in [1, 2]$
- テクスチャの複雑さ： $\mu \in [0, 1.5]$
- パターン類似度： $\nu \in [0, 1.5]$

次に、各20通りの修復結果の評価を行い、ラベル付けをした。この際、図22に示す評価用のインターフェースを作成して評価を行った。856枚に欠損の指定、修復を行い、評価によって515枚の品質について高のラベル付けがされた画像と663枚の品質について低のラベル付けがされた画像の計1,178枚の画像を含むファインチューニング用データセットを作成した。

4.1.2 畳み込みニューラルネットワークの学習

本実験は、深層学習のフレームワークであるChainer [37]を用いて行った。まず事前学習を行った。ここで、データセット204,512枚の内122,708枚を学習用に、40,902枚を評価用に、40,902枚をテスト用に分割して使用した。また、バッチサイズは20、エポック数は50回とし、Dropout Ratioは0.5とした。最適化手法にはAdam [38]を用い、Learning Rateを 2.0×10^{-4} とし、Weight Decay Rateを 1.0×10^{-4} とした。

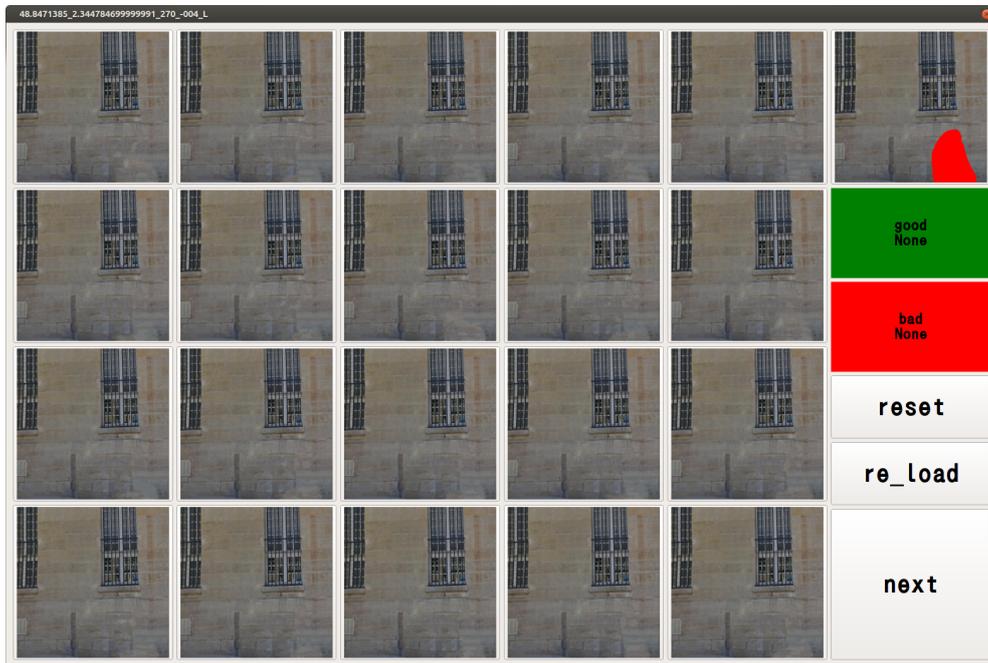


図 22: 評価に使用したインターフェース

次に事前学習での評価で一番誤差の小さかった 21 エポックでのネットワークと初期値とし、ファインチューニングを行った。データセット 1,178 枚の内 100 枚を評価用に、100 枚をテスト用に、残りを学習用に分割して使用した。また、バッチサイズは 5、エポック数は 100 回とし、dropout ratio は 0.5 とした。最適化手法には Adam [38] を使い、Learning Rate を 2.0×10^{-6} とし、Weight Decay Rate を 1.0×10^{-6} とした。評価において一番誤差の小さかった 98 エポックでのネットワークを最終的な学習結果とした。

学習したネットワークに、ファインチューニング用データセットのテスト用画像を入力し、識別率を求めた。また、転移学習の有用性を検証するため、事前学習のみを行いファインチューニングしない場合と、事前学習をせずファインチューニング用データセットのみを使用して学習する場合それぞれでも識別率を求め、提案手法と比較した。事前学習をしない場合では、事前学習をする場合と同じ方法でランダムな初期値から学習を行った。結果を表 1 にまとめる。ファインチューニングのみの場合が一番識別率が低く、転移学習を行った場合では、事前学習の

表 1: 識別率

	識別率 [%]
転移学習	59
事前学習のみ	57
ファインチューニングのみ	53

みの場合やファインチューニングのみの場合より識別率が向上しており，転移学習の効果がみられた．

4.2 学習済みネットワークによる修復画像の選出実験

学習済みネットワークによる修復画像の選出実験を行った．まず，4.1.1 節で述べたファインチューニング用のデータセット生成と同じ方法で，50 枚の画像に手動で欠損を与えた．これを 20 組のランダムなパラメータで修復した．次に，20 通りの修復画像を学習済みネットワークに入力し，品質に対するスコアを得る．20 種の内スコアの 1 番高いものを選出し，これをその画像に対しての提案手法の修復結果とした．

まず，修復画像とそのスコアを確認して，傾向を確認した．例として，元画像と欠損を与えた画像を図 23 に，20 通りの修復画像とそれぞれに対するスコアを図 24，図 25，図 26，図 27 に示す．ここで例 1，例 2 はパラメータによる修復結果の変化が大きい物，例 3 はパラメータによる修復結果の変化が小さくどの修復結果でも品質の悪いもの，例 4 はパラメータによる修復結果の変化が小さくどの修復結果でも品質の良いものである．

パラメータによる修復結果の変化が大きいものでは，図 24 のようにスコアに差がある場合と，図 25 のようにスコアの差が少ない場合があった．これは不自然なテクスチャの大きさ等の影響と考えられる．スコアに差がある場合では，修復結果の品質が良いもののスコアが高い傾向にあり，正しい選出結果を得ることができると考えられる．

また，パラメータによる修復結果の差があまりないものでは，スコアの差も小



(a) 例 1 : 元画像



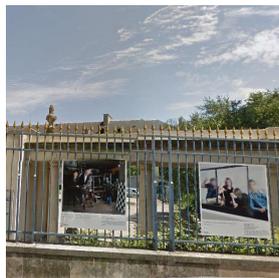
(b) 例 1 : 欠損を指定した画像



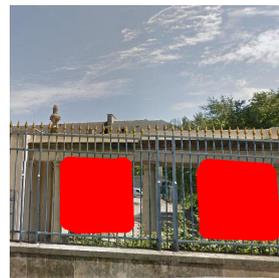
(c) 例 2 : 元画像



(d) 例 2 : 欠損を指定した画像



(e) 例 3 : 元画像



(f) 例 3 : 欠損を指定した画像



(g) 例 4 : 元画像



(h) 例 4 : 欠損を指定した画像

図 23: 欠損を与えた画像の例

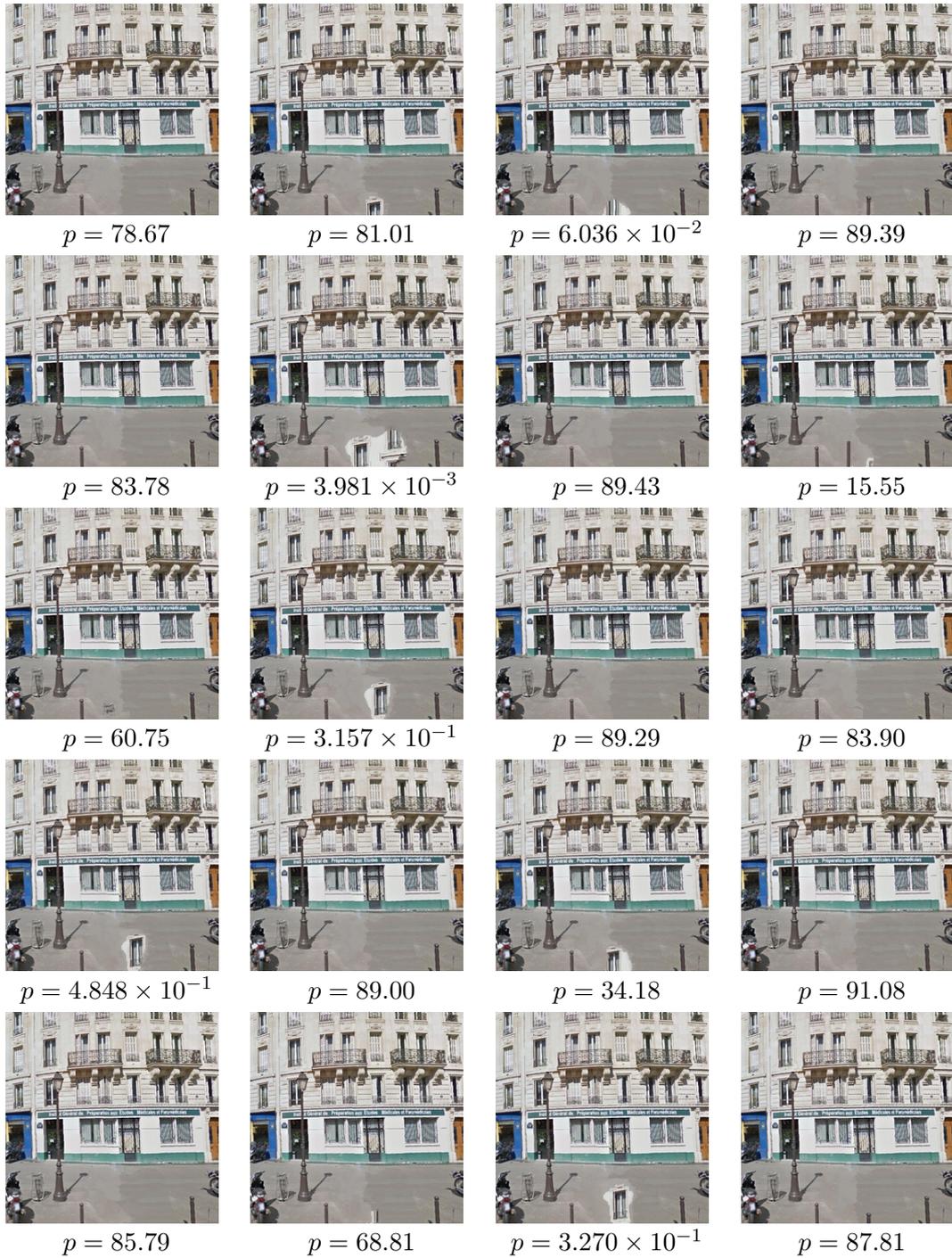


図 24: 修復画像の評価例 1 (p : 修復画像の品質についてのスコア)



図 25: 修復画像の評価例 2 (p : 修復画像の品質についてのスコア)

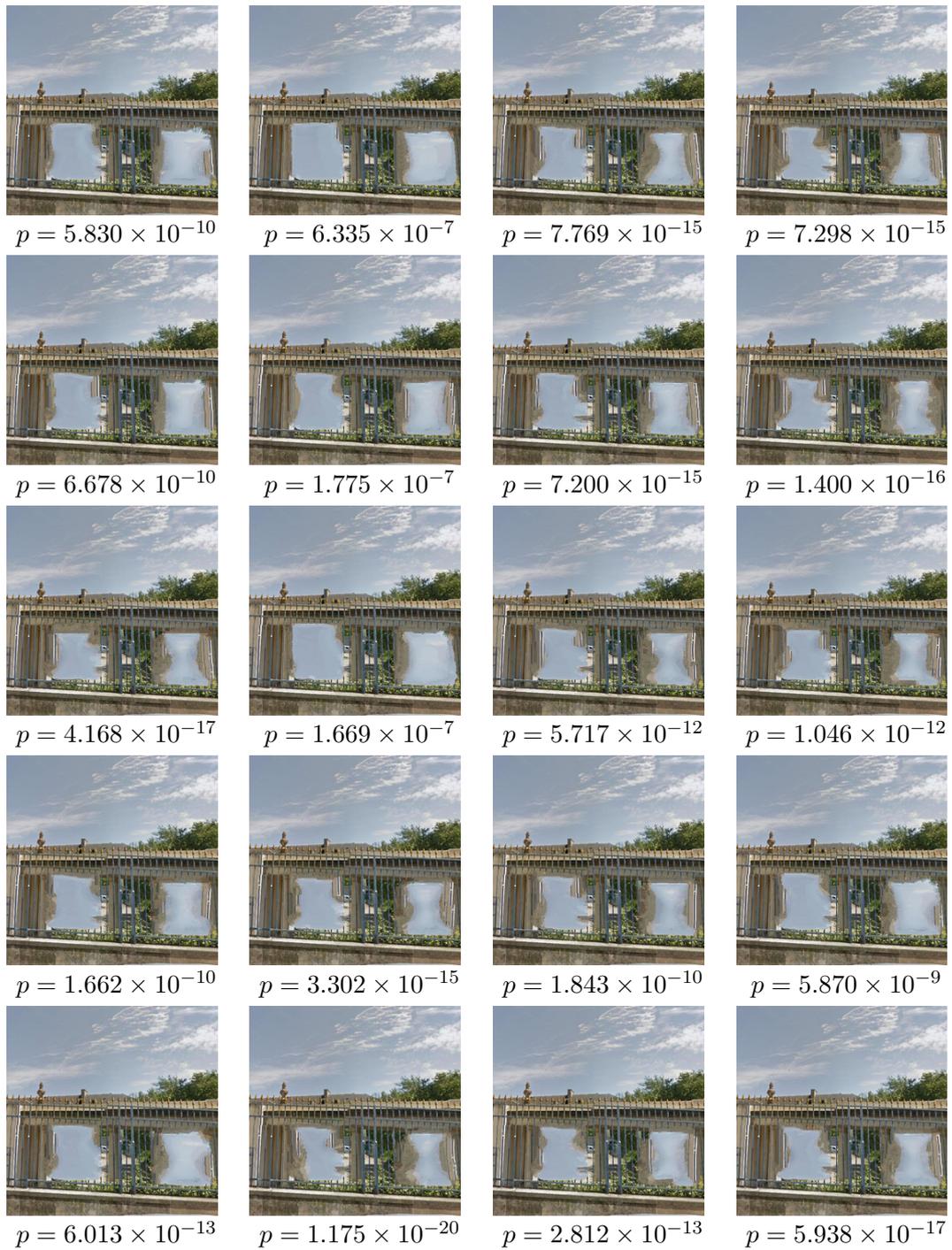


図 26: 修復画像の評価例 3 (p : 修復画像の品質についてのスコア)



図 27: 修復画像の評価例 4 (p : 修復画像の品質についてのスコア)



図 28: 評価に使用したインターフェース

さくなっていた。図 26 のような、どの結果も品質の悪いものではスコアが小さく、図 27 のような、どの結果も品質の良いものでは、スコアが大きくなっている。ここから、修復画像の品質とネットワークが出力したスコアは関連性があり、品質の高低がわかりやすい画像であれば識別できていると考えられる。

4.3 被験者による主観的評価実験

4.3.1 修復結果の手動による評価の検証

ファインチューニング用データセットの生成において、複数の修復結果の中から手動で品質の良いものと悪いものを選択しラベル付けを行う。この評価に個人差が無いかの確認実験を行った。4.1.1 節で生成した事前学習用データセットの中から、20 通りの修復結果の中に品質の良いものと悪いものの両方が存在した画像について、被験者に指定した欠損領域とそれぞれの画像を見せ、欠損領域に着目して品質の良いと思うものを選択させる。データセットの生成時に高のラベル付けをしたものが多く選択されれば、品質の評価に個人差が少ないと考えられる。実験では、図 28 に示すような評価用インターフェースを作成し、6 名の被験者に 50 組の画像を提示し、品質の良いと思うものを選択させた。

実験の結果、高のラベル付けをした画像が選択された確率は 88.67%であった。また、高のラベル付けをしたものを選択した人数毎の画像枚数を表 2 にまとめる。44 枚の画像で 5 人以上がラベルが高の画像を選択していることから、修復結果の手動による評価には個人差が少ないと考えられる。

表 2: ラベルが高の画像を選択した人数とその枚数

ラベルが高の画像を選択した人数 [人]	枚数 [枚]
0	0
1	2
2	1
3	2
4	1
5	12
6	32



(a) 欠損を指定した画像



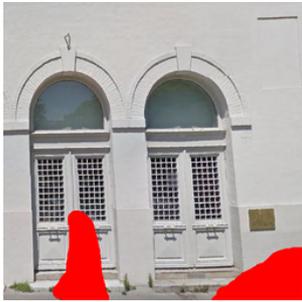
(b) ラベルが高の画像



(c) ラベルが低の画像

図 29: ラベルが高の画像を選択した人数が多い画像例 1

また、ラベルが高の画像を選択した人数が多い画像、少ない画像をそれぞれ確認し、傾向を調べた。ラベルが高の画像を選択した人数が多い画像の例として 6 人が選択した例を図 29、図 30 に、少ない画像の例として 1 人が選択した例を図 31 に、2 人が選択した例を図 32 に示す。ラベルが高の画像を選択した人数が多い画像は、図 29、図 30 で例に示したような 2 枚の修復結果が大きく異なるものが多い。こういった場合はラベル付けに個人差が少なくなると考えられる。ラベルが高の画像を選択した人数が少ない画像は、図 31、図 32 で例に示したような 2 枚の修復結果の差が小さいものも多く、修復した部分のどこに着目するか個人差が生じた結果、評価が一致しなかったと考えられる。



(a) 欠損を指定した画像

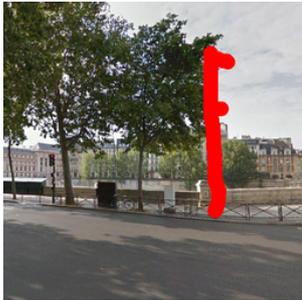


(b) ラベルが高の画像



(c) ラベルが低の画像

図 30: ラベルが高の画像を選択した人数が多い画像例 2



(a) 欠損を指定した画像



(b) ラベルが高の画像



(c) ラベルが低の画像

図 31: ラベルが高の画像を選択した人数が少ない画像例 1



(a) 欠損を指定した画像



(b) ラベルが高の画像



(c) ラベルが低の画像

図 32: ラベルが高の画像を選択した人数が少ない画像例 2

表 3: 提案手法による修復画像を選択した人数とその枚数

提案手法による修復画像を選択した人数 [人]	枚数 [枚]
0	6
1	7
2	6
3	3
4	14
5	3
6	11

4.3.2 学習済みネットワークによる修復画像の選出の検証

3.2 節で説明した，固定したパラメータによる画像修復の結果と，提案手法によって得られた画像修復の結果を比較する．4.2 節で述べた学習済みネットワークによる修復画像の選出実験で選出された画像と，同じ画像を 3 節で述べた，パラメータを固定して修復を行う従来手法を用いて修復した画像及び指定した欠損領域を見せ，欠損領域に着目して品質の良いと思うものを選択させる．選択された画像が提案手法によるものである確率を確認した．実験では，4.3.1 節での実験で用いたものと同じインターフェースを用いて，6 名の被験者に 50 組の画像を提示し選択させた．

実験の結果，提案手法によって選出された画像が選択された確率は 55%であった．また，提案手法による修復画像を選択した人数毎の画像枚数を表 3 にまとめる．提案手法による修復画像を選択した人数が多い画像，少ない画像，評価が割れた画像をそれぞれ確認し，傾向を調べた．

被験者全員が提案手法による修復画像を選択した例を図 33，図 34 に示す．図 33 の例では，従来手法による修復画像では不自然なテクスチャが存在したが，提案手法による修復画像では存在せず，図 34 の例では，提案手法による修復画像にも不自然なテクスチャが存在するが，違和感は軽減されている．また，品質についてのスコアも提案手法による修復画像が従来手法のものを上回っている．



(a) 欠損を指定した画像 (b) 従来手法による修復画像 (c) 提案手法による修復画像
 $p = 62.95$ $p = 64.09$

図 33: 被験者全員が提案手法による修復画像を選択した例 1 (p : 修復画像の品質についてのスコア)



(a) 欠損を指定した画像 (b) 従来手法による修復画像 (c) 提案手法による修復画像
 $p = 2.202 \times 10^{-13}$ $p = 2.125 \times 10^{-5}$

図 34: 被験者全員が提案手法による修復画像を選択した例 2 (p : 修復画像の品質についてのスコア)



(a) 欠損を指定した画像 (b) 従来手法による修復画像 (c) 提案手法による修復画像
 $p = 11.47$ $p = 76.43$

図 35: 被験者全員が従来手法による修復画像を選択した例 1 (p : 修復画像の品質についてのスコア)

被験者全員が従来手法による修復画像を選択した例を図 35, 図 36 に示す。提案手法による修復画像のスコアが、従来手法のものを上回っており、適切な評価が行えていない。またこれらの例では、パラメータを変化させて修復を行った画像の中に従来手法による修復画像と同程度の品質のものが存在した。

3人の被験者が提案手法による修復画像を選択した例を図 37, 図 38 に示す。これらの例では、パラメータの変化による修復結果の変化が小さい。この場合、2つの手法の差も小さくなりやすく、選択に個人差がみられたと考えられる。

以上から、ネットワークによる品質評価の精度が十分でなく適切な画像を選出できない場合があることや、パラメータを変化させて修復を行っても品質の良い修復画像を得られない場合があるため、今回のような結果になったと考えられる。被験者実験で提案手法による修復結果が選択されなかった画像でも、パラメータの違いによる複数の修復結果の中に選出された画像より品質の良い修復画像が存在する場合がある。このような場合であれば、ネットワークによる品質評価の精度を改善すれば、従来手法と同程度、またはより良い品質の修復画像を得ることができると考える。また、パラメータを変化させ修復を行っても品質の良い修復画像を得られない場合、主観による評価が難しく、評価の個人差があったと考えられる。



(a) 欠損を指定した画像 (b) 従来手法による修復画像 (c) 提案手法による修復画像
 $p = 8.400 \times 10^{-6}$ $p = 2.227$

図 36: 被験者全員が従来手法による修復画像を選択した例 2 (p : 修復画像の品質についてのスコア)



(a) 欠損を指定した画像 (b) 従来手法による修復画像 (c) 提案手法による修復画像
 $p = 51.01$ $p = 54.77$

図 37: 3人の被験者が提案手法による修復画像を選択した例 1 (p : 修復画像の品質についてのスコア)



(a) 欠損を指定した画像 (b) 従来手法による修復画像 (c) 提案手法による修復画像
 $p = 9.630 \times 10^{-1}$ $p = 12.27$

図 38: 3 人の被験者が提案手法による修復画像を選択した例 2 (p : 修復画像の品質についてのスコア)

5. まとめ

本論文では、画像修復においてより良い修復結果を得ることを目的とし、畳み込みニューラルネットワークを用いて、事例に基づいた手法による画像修復の結果から、自動的に品質の良いものを選出する手法を提案した。具体的には、必要となるデータセットを生成し、転移学習によって修復画像の品質を評価するための畳み込みニューラルネットワークを学習する。次に、修復したい画像をパラメータを変えながら複数回修復し、それらを学習済みネットワークに入力した上で、それぞれの品質の高低を評価して一番評価の高いものを選出する。

実験では、事前学習やファインチューニングの有無による品質の評価の精度を確認し、転移学習の有効性を示した。また、提案手法と従来手法を被験者による主観的評価実験によって比較し、修復結果の品質に向上が見られたかを検証した。その結果、提案手法によって品質が向上した場合とそうでない場合が存在したため、ネットワークの識別性能の向上が必要であることがわかった。パラメータの変化によって修復結果に大きな違いがある画像であれば、識別性能の向上によって品質の向上が見込める。

今後は、品質評価の性能の向上のため、データセットのラベルの精度の向上やデータ数の増加を行う。また、ネットワークの構造や学習率などのパラメータを調整する。また、提案手法で利用していないデータセットや画像修復手法においても提案手法が適用可能かを検証する。

謝辞

本研究を進めるにあたり，懇切なる御指導，御鞭撻を頂いた視覚情報メディア研究室 横矢直和 教授に心より感謝申し上げます。また，本研究の遂行にあたり，有益な御助言，御鞭撻を頂いた光メディアインタフェース研究室 向川康博 教授に厚く御礼申し上げます。そして，本研究を進めるにあたり，細やかな御指導，御鞭撻を頂いた視覚情報メディア研究室 佐藤智和 准教授に深く感謝致します。また，本研究に対して終始的確な御指導，御鞭撻を頂いた視覚情報メディア研究室 中島悠太 客員准教授に深く御礼申し上げます。さらに，本研究を行うにあたって多大なる御助言，御鞭撻を頂いた視覚情報メディア研究室 河合紀彦 助教に心より感謝します。また，研究室での生活を支えていただいた視覚情報メディア研究室秘書 石谷由美女史，南あずさ女史に厚く御礼申し上げます。そして，研究だけでなく，日々の研究室生活においても大変お世話になりました視覚情報メディア研究室の皆様にも深く感謝します。最後に，両親をはじめ，大学院生活を支えてくださったすべての方々に感謝の意を表します。

参考文献

- [1] 前田浩幸, 高橋健一, 太田正光. 欠損画像の修復処理の一方式. 電子情報通信学会論文誌 D, Vol. J69-D, No. 1, pp. 91–97, 1986.
- [2] Simon Masnou and Jean-Michel Morel. Level lines based disocclusion. In *Proc. IEEE Int. Conf. on Image Processing*, pp. 259–263, 1998.
- [3] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. Image inpainting. In *Proc. ACM SIGGRAPH*, pp. 417–424, 2000.
- [4] Jianhong Shen, Sung H. Kang, and Tony F. Chan. Euler’s elastica and curvature-based inpainting. *SIAM Journal on Applied Mathematics*, Vol. 63, No. 2, pp. 564–592, 2003.
- [5] 小川貴弘, 長谷山美紀, 北島秀夫. オプティカルフローを用いた静止画像における失われた輝度値の復元. 電子情報通信学会論文誌 D2, Vol. J87-D-2, No. 9, pp. 1786–1795, 2004.
- [6] 東海林健二. テクスチャ画像における欠損部修復の一手法. 電子情報通信学会論文誌 D, Vol. J71-D, No. 9, pp. 1701–1708, 1988.
- [7] Anil N. Hirani and Takashi Totsuka. Combining frequency and spatial domain information for fast interactive image noise removal. In *Proc. ACM SIGGRAPH*, pp. 269–276, 1996.
- [8] 天野敏之, 佐藤幸男. 固有空間法を用いた BPLP による画像補間. 電子情報通信学会論文誌 D2, Vol. J85-D-2, No. 3, pp. 457–465, 2002.
- [9] 天野敏之, 佐藤幸男. kBPLP 法を用いた高次元非線形射影による画像補間. 電子情報通信学会論文誌 D2, Vol. J86-D-2, No. 4, pp. 525–534, 2003.
- [10] 井添慎太郎, 剣持雪子, 小谷一孔. 画像のフラクタル性に基づいたカーネル主成分分析による遮へい領域の原画像推定. 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, Vol. 101, No. 362, pp. 55–62, 2001.

- [11] Alexei A. Efros and Thomas K. Leung. Texture synthesis by non-parametric sampling. In *Proc. IEEE Int. Conf. on Computer Vision*, Vol. 2, pp. 1033–1038, 1999.
- [12] Marcelo Bertalmio, Luminita Vese, Guillermo Sapiro, and Stanley Osher. Simultaneous structure and texture image inpainting. *IEEE Trans. on Image Processing*, Vol. 12, No. 8, pp. 882–889, 2003.
- [13] Paul Harrison. A non-hierarchical procedure for re-synthesis of complex textures. In *Proc. Int. Conf. in Central Europe Computer Graphics, Visualization and Computer Vision*, pp. 190–197, 2001.
- [14] Iddo Drori, Daniel Cohen-Or, and Hezy Yeshurun. Fragment-based image completion. *ACM Trans. on Graphics*, Vol. 22, No. 3, pp. 303–312, 2003.
- [15] Raphaël Bornard, Emmanuelle Lecan, Louis Laborelli, and Jean-Hugues Chenot. Missing data correction in still images and image sequences. In *Proc. ACM Int. Conf. on Multimedia*, pp. 355–361, 2002.
- [16] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. on Image Processing*, Vol. 13, No. 9, pp. 1200–1212, 2004.
- [17] Wen-Huang Cheng, Chun-Wei Hsieh, Sheng-Kai Lin, Chia-Wei Wang, and Ja-Ling Wu. Robust algorithm for exemplar-based image inpainting. In *Proc. Int. Conf. Computer Graphics, Imaging and Visualization*, pp. 64–69, 2005.
- [18] Sebastian Vetter, Marcin Grzegorzec, and Dietrich Paulus. Wavelet-based inpainting for object removal from image series. In *Joint Pattern Recognition Symp.*, pp. 343–352, 2010.
- [19] BianRu Li, Yue Qi, and XuKun Shen. An image inpainting method. In *Proc. IEEE Int. Conf. Computer Aided Design and Computer Graphics*, pp.

- 531–536, 2005.
- [20] Zongben Xu and Jian Sun. Image inpainting by patch propagation using patch sparsity. *IEEE Trans. on Image Processing*, Vol. 19, No. 5, pp. 1153–1165, 2010.
- [21] Nikos Komodakis. Image completion using global optimization. In *Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 442–452, 2006.
- [22] Yael Pritch, Eitam Kav-Venaki, and Shmuel Peleg. Shift-map image editing. In *Proc. IEEE Int. Conf. on Computer Vision*, Vol. 9, pp. 151–158, 2009.
- [23] Cédric Allène and Nikos Paragios. Image renaissance using discrete optimization. In *Proc. IAPR Int. Conf. on Pattern Recognition*, Vol. 3, pp. 631–634, 2006.
- [24] Yonatan Wexler, Eli Shechtman, and Michal Irani. Space-time completion of video. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 3, pp. 463–476, 2007.
- [25] Soheil Darabi, Eli Shechtman, Connelly Barnes, Dan B. Goldman, and Pradeep Sen. Image Melding: Combining inconsistent images using patch-based synthesis. *ACM Trans. on Graphics*, Vol. 31, No. 4, pp. 82:1–82:10, 2012.
- [26] Norihiko Kawai, Tomokazu Sato, and Naokazu Yokoya. Image inpainting considering brightness change and spatial locality of textures and its evaluation. In *Proc. Pacific-Rim Symp. on Image and Video Technology*, pp. 271–282, 2009.
- [27] Norihiko Kawai and Naokazu Yokoya. Image inpainting considering symmetric patterns. In *Proc. IAPR Int. Conf. on Pattern Recognition*, pp. 2744–2747, 2012.

- [28] Jia-Bin Huang, Sing B. Kang, Narendra Ahuja, and Johannes Kopf. Image completion using planar structure guidance. *ACM Trans. on Graphics*, Vol. 33, No. 4, p. 129, 2014.
- [29] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Trans. on Graphics*, Vol. 28, No. 3, p. 24, 2009.
- [30] Junyuan Xie, Linli Xu, and Enhong Chen. Image denoising and inpainting with deep neural networks. In *Proc. Neural Information Processing Systems*, pp. 341–349, 2012.
- [31] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.
- [32] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A Efros. Context encoders: Feature learning by inpainting. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2536–2544, 2016.
- [33] Sinno J. Pan and Qiang Yang. A survey on transfer learning. *IEEE Trans. on Knowledge and Data Engineering*, Vol. 22, No. 10, pp. 1345–1359, Oct 2010.
- [34] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proc. ICML Int. Conf. on Machine Learning*, pp. 448–456, 2015.
- [35] Nitish Srivastava, Geoffrey E. Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, Vol. 15, No. 1, pp. 1929–1958, 2014.

- [36] Carl Doersch, Saurabh Singh, Abhinav Gupta, Josef Sivic, and Alexei A. Efros. What makes paris look like paris? *ACM Trans. on Graphics*, Vol. 31, No. 4, pp. 101:1–101:9, 2012.
- [37] Seiya Tokui, Kenta Oono, Shohei Hido, and Justin Clayton. Chainer: a next-generation open source framework for deep learning. In *Proc. Neural Information Processing Systems*, 2015.
- [38] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.