



会議レポート

CVPR 2018 参加報告

国際会議 CVPR の概要

CVPR は、IEEE Computer Society と Computer Vision Foundation (CVF) 主催による、コンピュータビジョン (CV) とパターン認識 (PR) に関する国際会議 (IEEE/CVF Conference on Computer Vision and Pattern Recognition) である。ICCV や ECCV と同等の難易度を持つ CV 系トップ会議の 1 つであり、毎年 6 月中旬ごろに米国で開催され、今年で 31 回目を迎える。論文は IEEE Xplore ほか、CVF によるオープンアクセス公開もある¹⁾。

今年は 2018 年 6 月 18 日から 22 日にかけてソルトレイクシティで開催され、去年の記録を塗り替える 6,500 人が参加するなど大変盛況であった。論文の投稿数も記録を更新し (図-1)、今年は 3,359 本の投稿のうち、例年通り 29% にあたる 979 本が採択された。エリアチェアが 108 人、査読者が 1,715 人という大規模な査読体制の下、1 本の論文に 3 名の査読者と最低 3 名のエリアチェアがつき採否が決まる。公平性のため、プログラムチェアは投稿が許されていない。

本会議の 3 日間では、全論文がポスターで発表されるほか、一部の論文は口頭発表の舞台も用意され、Facebook のページからは生中継が見られる。会議の規模拡大に伴い、ポスターセッションは 10 時～12 時半、12 時半～15 時、16 時半～18 時半などの枠が設けられ、うち昼の 2 時間は提供される食事を自由にとることができた。オーラルセッションは昨年より 3 トラックの並列になり、1 セッションに 12 分 + 質疑 2 分の発表が 4 件ほどと、さらにスポットライトという質疑なしの 4 分の発表が 10 件ほど続く。オーラルの採択率は 2.1%、スポッ

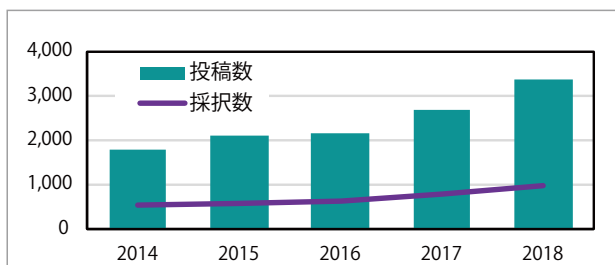


図-1 過去 5 年の CVPR の投稿数と採択数

トライトは 6.6% と、誉れ高い舞台である一方、ポスター発表のみの論文も同等の業績として記録される。

前後の 2 日間ではチュートリアルが 21 件、ワークショップが 49 件開催された。チュートリアルでは自動運転、敵対的生成ネットワーク (Generative Adversarial Nets : GAN)^{☆1}、弱教師あり / 教師なし学習が多く、またワークショップの多くは、開催者が用意した学習データで競わせるコンペ形式だった。本会議中は企業のデモ展示もあり、ブースには米国 IT 大手をはじめ、急成長する中国企業など 140 社が並び、内容は自動運転系が多く、その次に AI ソリューション、ロボット、フィンテック、e コマース、仮想現実 (VR) と続いた。会場の様子を 図-2, 3 に示す。その他の工夫として、会議専用アプリにより日程の確認や参加者との連絡ができたほか、ワールドカップの中継スクリーン (図-4) や休憩用の卓球台が設置されていた。夜はレセプションのみでバンケットはなかったが、代わりにたくさんの企業からパーティの招待があるようで、夕食には困らなかった。

☆1 敵対的生成ネットワーク (GAN) とは、何らかの確率分布 (たとえば一様乱数) を入力とし特殊な偏りを持つ分布 (たとえば画像) を生成する生成器 (generator) と、生成された分布と目標の分布 (画像) を比較し、生成されたものがどうかを判定する識別器 (discriminator) を用意し、生成器は識別器を騙すように、識別器は騙されないように、それぞれ最適化させると、最終的に生成される分布は目標とする分布に必ず近づく、という事実に基づき、所望の分布を生成させるネットワークのこと。生成器と識別器が互いを負かすように学習を行うため「敵対的」という名を冠する。2014 年に提案され、本物の写真と区別できないような画像が生成できるため大変話題になった。



図-2 メイン会場の 1 つ



図-3 デモとポスターブース

近年の CVPR の研究動向

年々、研究動向の変化が加速しているが、今回改めて近年の激しい変化を感じた。そこで、主に過去5年のCVPRについて、動向を振り返りたい。

CVPR 2014～2016は、深層学習（ディープラーニング）がCV分野に浸透するまでの過渡期であった。特にCVPR 2014では、2012年末のAlexNet（Krizhevskyら、2012）の登場から加熱した深層学習による画像認識の研究が、物体検出（GirshickらのR-CNN）や顔認識（TaigmanらのDeepFace）などの周辺タスクへ飛び火した。一方で、このころはまだ深層学習を用いない従来研究も盛んで、事実、ベストペーパーを始めとする3本の受賞論文は、光の反射や数理最適化に関する研究などであった。

CVPR 2015には、深層学習が本格的に台頭しはじめ、深層学習を含む研究の割合が初めて過半数を占めた（筆者調べ）。画像認識ではネットワークの深層化が進み（Szegedyらの22層GoogLeNet）、また、ステレオ^{☆2}などのlow-levelなタスクにも深層学習が適用され始めたが、この時点では従来手法の一部を深層学習で置き換えるにとどまっていた。これを打開したのは、おそらく受賞論文にも選ばれたLongらの全層畳み込みネットワーク（Fully Convolutional Networks: FCN）で、これ以降、画像分割やステレオなど任意サイズの画像入出力を扱うさまざまなタスクに深層学習が適用可能になった。一方、深層学習以外の研究もまだ存在感があり、合計5本の受賞論文のうち、ベストペーパー（Newcombeらの非剛体3次元形状推定DynamicFusion）を含む3本は、深層学習以外の研究だった。

CVPR 2016は、深層学習がほぼ完全にCVPRに浸透し、ステレオなど多くのCVタスクでEnd-to-end学習が実現した（MayerらのDispNetやFlowNetなど）ほか、いくつかの重要な技術や概念が登場した時期でもあった。たとえば、ベストペーパーであるHeらのResNetや、

☆2 ステレオとは、2枚以上の画像とその画像を撮影したカメラの位置や姿勢から、画像内の3次元を復元する技術のこと。人間も両眼のステレオ視で立体視をしている。



図-4 サッカーワールドカップを観戦する人々

前年に発表されたバッチ正規化（Ioffe&Szegedy, ICML 2015）は、深層学習の必須技術として広く浸透した。GAN関連の研究も登場し始め、特にPathakらはGANを敵対損失（adversarial loss）^{☆3}として応用し、画像生成以外のタスクに適用させた。また、GatysらのNeural Style Transfer^{☆4}もここで発表された。一方で、従来研究は存在感を減らし、全体として8割以上が深層学習を含む研究だったと思う。

こうして2017年には、FCNやResNet、バッチ正規化、GAN/敵対損失など、深層学習において今や当然となった技術が出揃い、学習データさえあれば大抵のタスクは比較的容易に学習できる環境が整った。そうした状況のCVPR 2017では、性能向上を目指したネットワーク構成の工夫に関心が高く、ResNetの発展系（HuangらのベストペーパーDenseNet）のほかに、ステレオ等における古典的アプローチを深層ネットワークで模倣する試みも散見された（Ranjan&BlackのSPyNetなど）。また、強化学習を深層学習に適用する深層強化学習も目立った。

今年の CVPR の研究動向

今年のCVPRでは、ImageNetタスクに代表されるような、正解ラベル付き学習データが存在する前提での、ネットワーク構成の工夫による性能向上の研究は、ひとまず収束したと感じた。代わりに、学習データがない（少ない）状況での学習方法や、従来研究の見聞やシーンの生成過程に関する事前知識を深層学習に取り入れる方法などの議題が目立った。これらについて、キーワードとともに紹介したい。

1) 3D Vision, 教師なし学習, Disentangling

今年のCVPRは3D Visionの研究が大変豊作だった。3D Visionは元々コスト関数の最小化による学習なしの手法が主流なため、実はデータ・ハングリーな深層学習との相性が良い。たとえばステレオでは、輝度残差に基づき推定モデルと観測画像との一致度を評価するが、これを深層学習で再構成損失として用い、教師なし学習を実現する事例が今年多く見られた。これは見方を変えれば、従来のように単にネットワークに正解出力を与えて覚えさせるのではなく、入力データ（観測画像）が生成された原理（3次元幾何や投影モデル）や構成要素（シーン形状、物体・カメラモーション等）をきちんと考え、ネットワーク上で生成過程を再現するようになったといえる。このように生成過程を紐解き、構成要素に分解して再構

☆3 Pathakらは、GANで提案された損失関数（loss function）を損失関数の1つとみなし、画像の再構成誤差と組み合わせた。多くの論文がその後同様にadversarial lossを別の損失関数と組み合わせ、さまざまなタスクに利用している。

☆4 写真をピカソやゴッホが描いたように変換できることで話題になった。この技術を使ったPrismaというアプリが無料で利用できる。

成するアプローチは Disentangling と呼ばれ、厳格な物理モデルが存在する 3D Vision との相性が良いこともあり、今 CVPR の 1 つのキーワードになっていた。

2) 強化学習, ドメイン適応

学習データ不足という問題意識のもと、強化学習とドメイン適応^{☆5}の研究は非常に盛んであった。特に強化学習は、CVPR の 2 週間後にスウェーデンで行われた機械学習分野のトップ国際会議 ICML 2018 において、最も盛んなトピックとして実に全体の論文の 10% を占めていた。ドメイン適応の解法では GAN の応用が多かった。

3) 不規則データ構造に対する畳込み^{☆6}

点群やグラフといった、グリッド以外のデータ構造に対してどのように畳込みを定義するかという問題は、残された大きな課題の 1 つとして、今回も 10 例ほどの研究が発表された。特に Su²⁾ の SPLATNet では、3D 点群データを特殊なメッシュ平面にマッピングして畳み込む手法を提案し、受賞論文に選ばれた。このような演算により 3D 点群の認識が高精度かつ高速に行える。

4) その他の頻出技術

敵対損失は、超解像や単眼デプス推定などさまざまな

☆5 ドメイン適応とは、ソースドメインでの教師データと、ターゲットドメインでの観測値を与えられた状態で、ターゲットドメインでのタスクの学習を目指す問題である。

☆6 畳込みとは、ある関数に、その定義域よりも小さな定義域を持つ関数を少しずつずらしながら積分を計算していく演算のこと。画像のフィルタリングによく用いられる。

タスクに取り入れられていた。また、従来手法に使われていた最適化手法の反復処理を、展開 (unroll) して微分可能なネットワーク層として組み込む方法も、確立したテクニックとして散見された。たとえば、Krishna³⁾ はメッセージ伝搬法をネットワークに組み込むことで隣接物体の関係性を利用したり、また、受賞論文でもある Zafir⁴⁾ の研究では、グラフ・マッチングの最適化処理を組み込むことで、特徴量からマッチングまでを End-to-end で学習する手法を提案している。

CVPR 2019

2019 年の CVPR はカリフォルニアのロングビーチで 6 月 15 日～21 日に開催される。論文締切は 2018 年 11 月 16 日である。弱教師あり/教師なし学習、強化学習、ドメイン適応、音声や自然言語処理と融合した高度な「意味」知識の獲得、マルチメディア処理、ロボットへの画像による教示、などが多く提案されることだろう。

参考文献

- 1) CVF Open Access, <http://openaccess.thecvf.com/menu.py>
- 2) Su, H. et al. : SPLATNet : Sparse Lattice Networks for Point Cloud Processing, CVPR18.
- 3) Krishna, R. et al. : Referring Relationships, CVPR18.
- 4) Zafir, A. and Sminchisescu, C. : Deep Learning of Graph Matching, CVPR18.

(谷合竜典/理化学研究所 革新知能統合研究センター、川上 玲/東京大学情報理工学系研究科)

書評 (ビブリオ・トーク)・会議レポート募集のお知らせ

情報処理学会会誌編集委員会では、会誌「情報処理」に掲載する書評、および会議レポートを広く会員の皆さまから募集しています。

1. 募集対象 次の 2 種類の記事について、原稿を募集します。書評に関しては、「ビブリオ・トーク—書評—」、「ビブリオ・トーク—私のオススメ—」の 2 つのカテゴリを設けます。

a-1) ビブリオ・トーク—書評—: 過去 2 年間に出版された、本会会員にとって有益な図書についての紹介もしくは批評。

a-2) ビブリオ・トーク—私のオススメ—: お気に入りの本の紹介。

b) 会議レポート: 情報処理に関する国際規模の会議・大会の報告など、時事性が高く、本会会員に広く知らせる価値のある話題。

2. 応募資格

原則として本会会員に限ります。

3. 応募の手続き

1) 表 題: ビブリオ・トークの場合は、書評もしくは私のオススメの投稿カテゴリ、著者名、書名、ページ数、発行所、発行年、価格、ISBN を書く。会議レポートは、見出しを書く。書評、会議レポートの別を左肩に書く。

2) 評者名 (会議レポートの場合は筆者名)・所属・評者連絡先 (住所、E-mail、Fax など) の記載を忘れずに。

3) 本 文: ビブリオ・トークは 1,500 字以内または 3,000 字以内 (1 または 2 ページ)。会議レポートは 2,100 字前後で書く。

4) そ の 他: (必要であれば) 参考文献、付録、図、表をつける。詳しくは「原稿執筆のご案内 / 書評・会議レポート」(<https://www.ipsj.or.jp/magazine/sippitsu/shohyonews.html>) を参照してください。

4. 原稿の取扱い

投稿された原稿は会誌編集委員会で審査し、採否を決定します。採用にあたっては原稿の修正をお願いすることがあります。あらかじめご了承ください。

5. 照会/応募先 一般社団法人 情報処理学会 会誌編集部門 E-mail: editj@ipsj.or.jp

