

Twitter画像に対する地域別画像タイプの大規模分析

長野 哲也, 會下 拓実, 柳井 啓司 (電気通信大学)



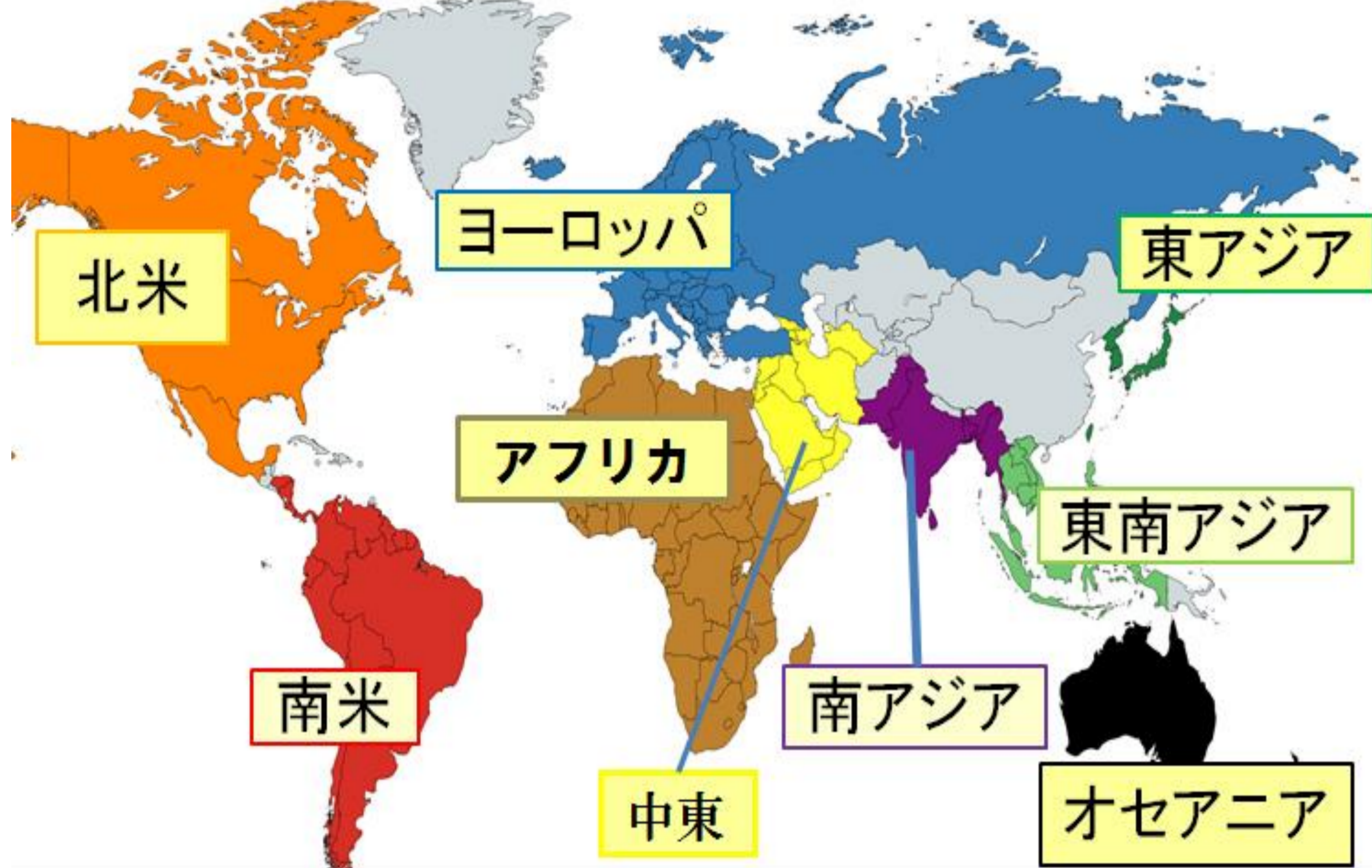
PS2-36

1. 背景&目的

Facebook、Instagram、TwitterなどのSNS上に世界各地で撮られた画像が大量に存在



大量のTwitter画像を画像のジャンルで分類し、地域別で分析し、地域別の画像の特徴を発見。分析の際には、テキストを解析せずに分析。



本研究における分析地域

2. 手法

・画像収集

位置及び投稿時間、画像付きツイートログを使用し画像を収集

・画像分類

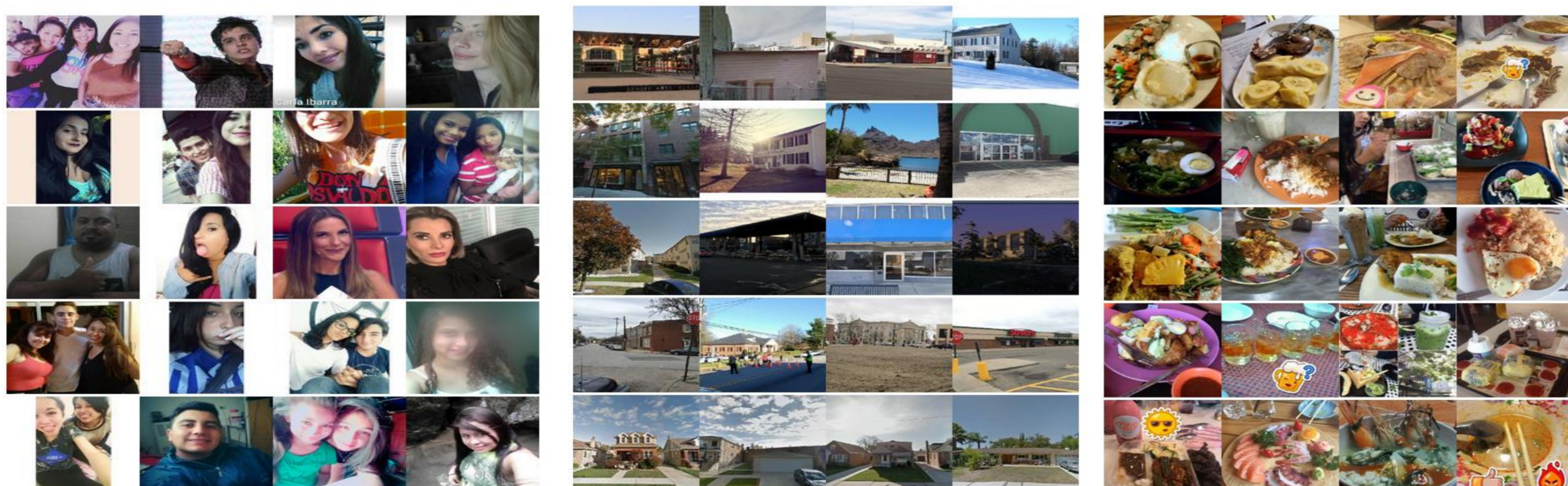


Deep Learningで画像の特徴量を抽出。Deep Learningの使用により、分類の精度が飛躍的に向上。特徴量を用いてK-meansで画像を分類

・画像分析

分類結果から、画像のジャンルを分析し、分析地域共通の画像ジャンルを検出。ジャンルは「人」「建物」「食べ物」「風景」「文章」の5ジャンルに集約

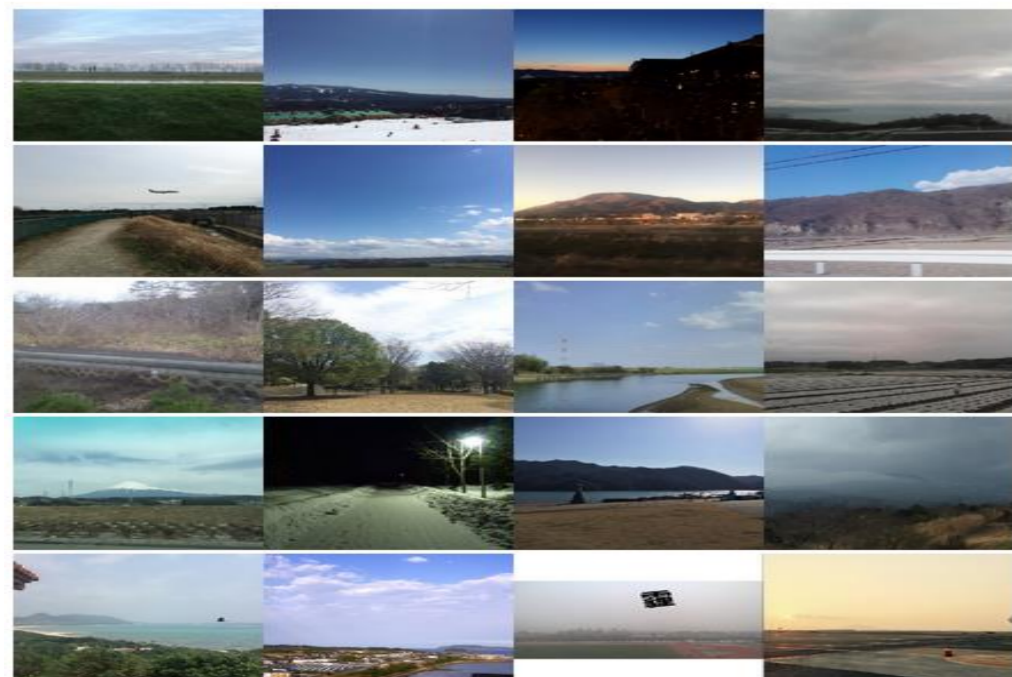
地域別の画像ジャンルの分布割合から各地域の特性を分析。



人の画像

建物の画像

食べ物の画像



風景の画像



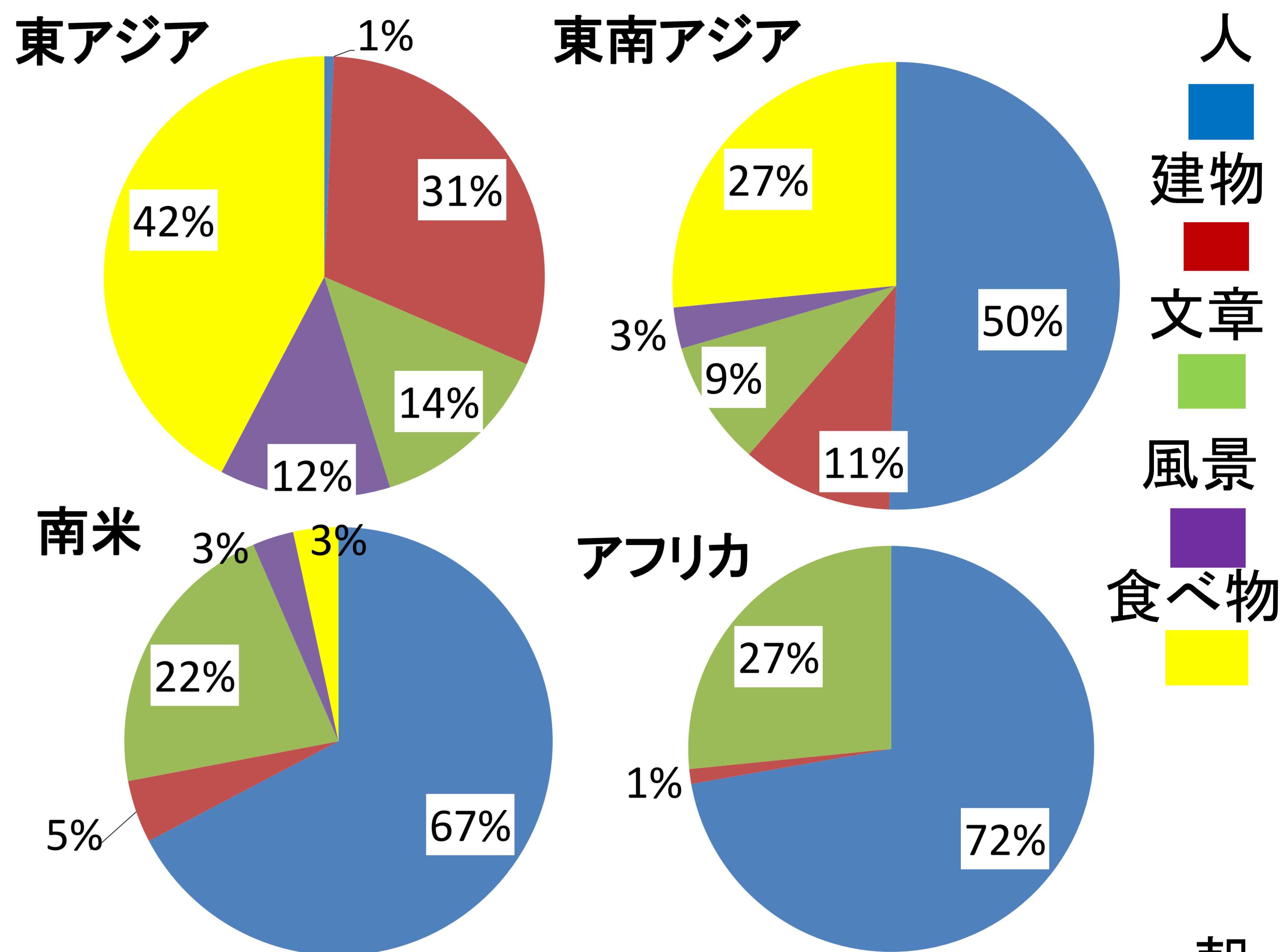
文章の画像

4. まとめと今後の課題

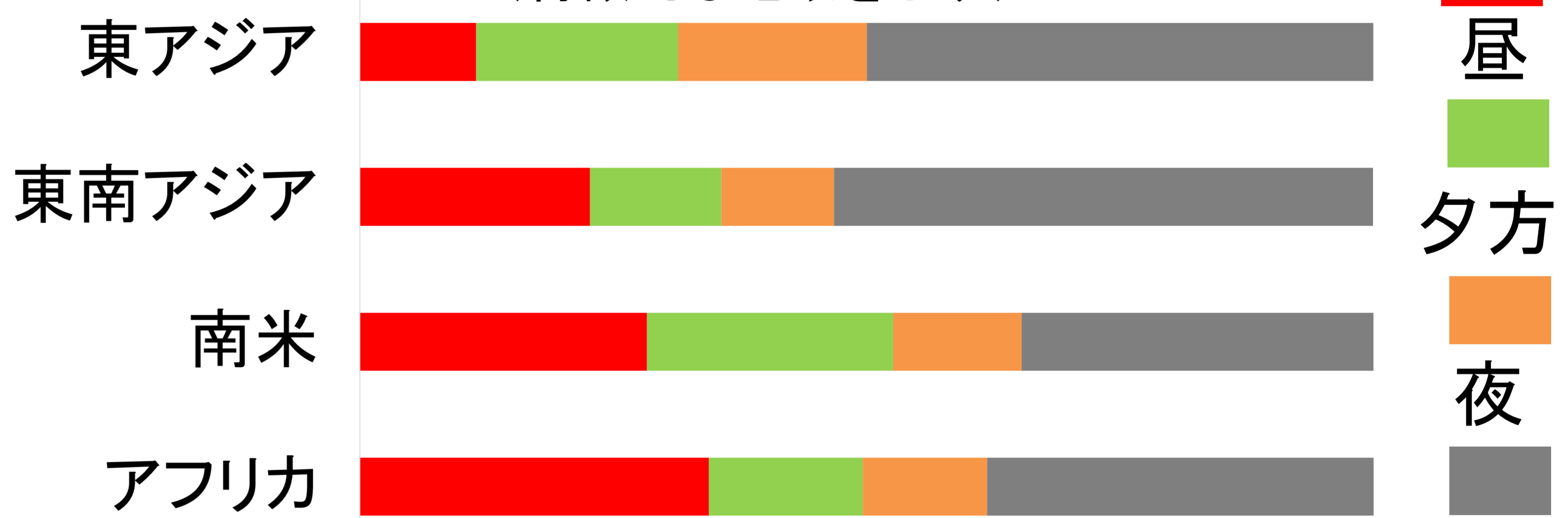
- ・ テキストを使用せず、Twitter画像と位置情報と投稿時間のデータから地域別の特徴を検出。
- ・ 分析結果から地域別の画像の傾向や投稿の時間帯の傾向を発見
- ・ 本実験で定めた5種類の共通画像のジャンルの細分化により、新たな地域の特徴を発見

3. 実験

2016年1月から6月までのTwitterログから画像を約216万枚収集し、地域ごとの画像分類結果を人手で確認。画像を5ジャンル+該当なしに分類し、ノイズ除去後、各地域の5ジャンルの画像を集計し、割合を出力、各地域の特徴を分析



5ジャンルの画像の割合を示す円グラフ (特徴的な地域を示す)



1月から6月までの地域別の時間帯の投稿割合

・地域別分析

- ・ 東アジアは人の画像が少なく、食べ物の画像が多い。
- ・ 東南アジアは人の画像と食べ物の画像が多い。
- ・ 南米とアフリカは人の画像の割合が約7割ととても多く、残りは文章の画像が約2割で他は少なかった。
- ・ グラフ以外の地域では、南アジア、オセアニアは文章の画像が多く、北米、ヨーロッパ、中東では5つのバランスが取れている。

・時間帯分析

- ・ 東アジアは朝の投稿が少なく、夜の投稿が多い。
- ・ 南米は朝昼夜と投稿されていた。
- ・ 他の地域では、北米が南米と同様の傾向、ヨーロッパ、アフリカ、中東、南アジア、オセアニアは朝と夜の投稿が多い結果となった。