

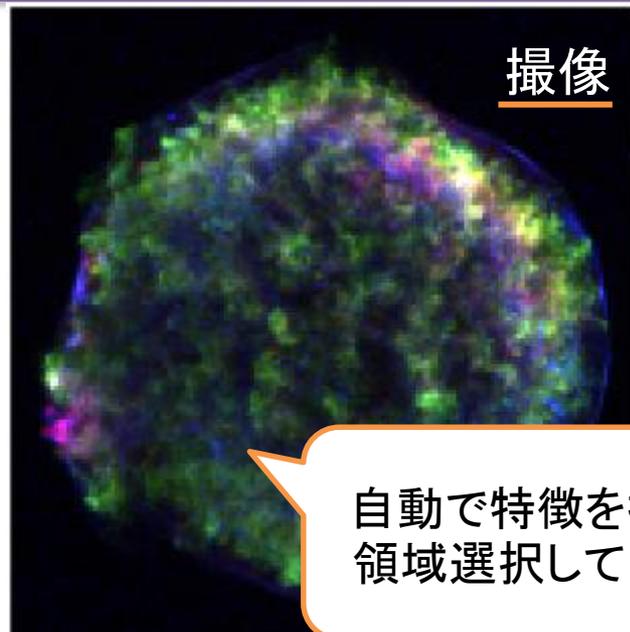
教師なし深層学習を用いた超新星残骸のX線分光研究

岩崎啓克（立教大学）

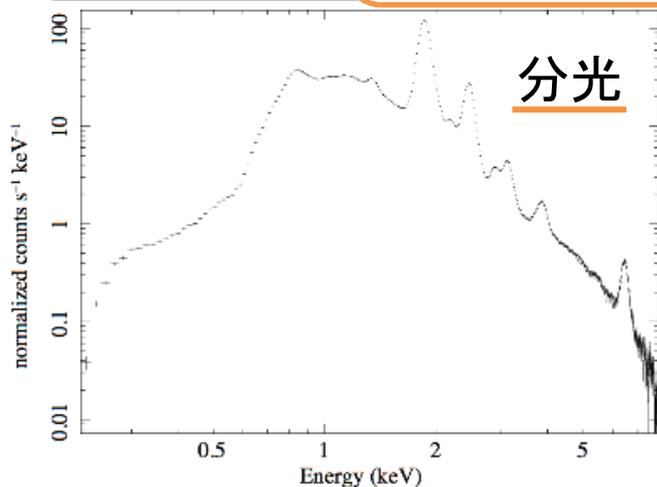
2020年3月2日

高宇連

超新星残骸への機械学習応用のモチベーション



自動で特徴を捉えて
領域選択してほしい



X線観測は同時に撮像 & 分光
空間と光子エネルギーの情報を持った
複雑な高次元データが得られる

さらに将来の観測機計画では分解能が
圧倒的に向上する(150 eV \rightarrow 2.5 eV
@6 keV \times 約4000 pix, Athena)

現状では、人の目で特徴を把握し、
人の手で領域選択する



複雑な観測データから機械的な特徴
抽出をしたい。人のバイアスも極力省く



機械学習を利用

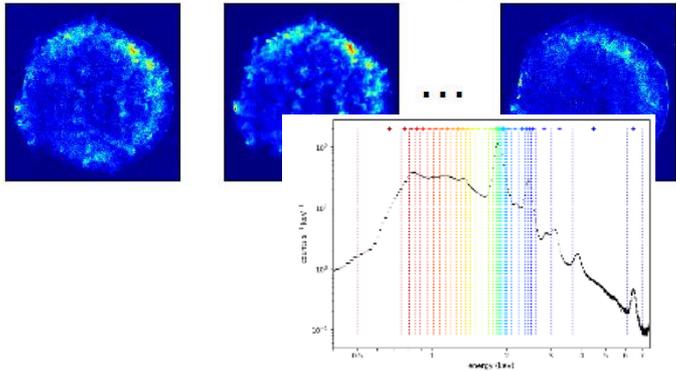
天文観測への機械学習の応用の状況

- ✓ 既にビッグデータ化している可視光観測データに対する応用、特に教師あり分類は進んでいる。深層学習も利用されている。e.g., 銀河、超新星の分類
- ✓ 機械学習の応用がまだあまり進んでいない分野も多い。
 - 天体個々の詳細解析; 時間変動や内部構造の解析など
 - 観測データから教師なし学習の手法での特徴抽出
- ✓ 超新星残骸(SNR)の中では物理状態・放射過程が場所によって異なる。空間分解したスペクトル解析が必要。
- ✓ これまでSNR観測に対しては、主成分分析(PCA)やクラスタリングの線形な手法の適用のみ
 - PCAで幅の広いエネルギービンのデータを次元削減 (Warren+05)
 - PCAでSi K α 輝線のシフトを抽出 (Sato & Hughes 17)
 - 幅の広いエネルギービンに対しクラスタリング(Burkey+13)

VAEとGMMを組み合わせた教師なし分類の新手法

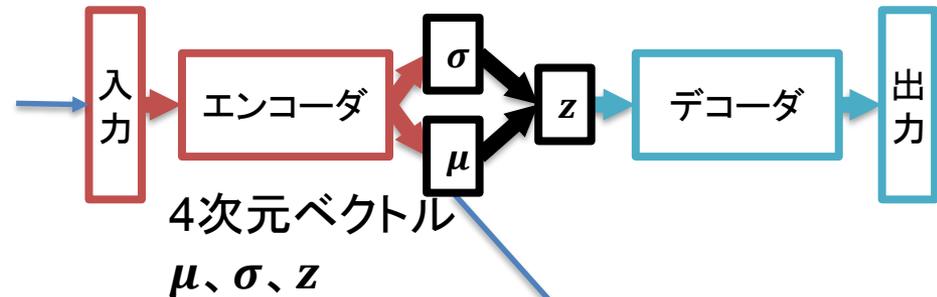
変分オートエンコーダ(VAE)と、混合ガウスモデル(GMM)を組み合わせ、**非線形な高次元データの自動分類**を考案。スペクトル情報のみを用い、空間構造を抽出。

① 天体全体のスペクトルを自動で37分割。スペクトル(37次元ベクトル)のデータセットを用意。



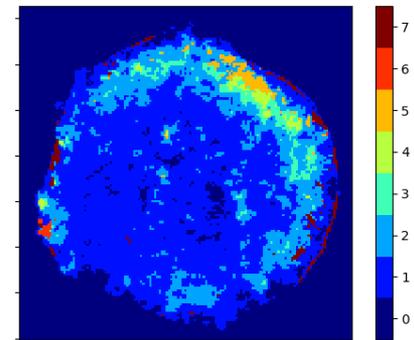
37次元ベクトル x

② 各空間ビンのスペクトルをVAEで次元削減。4次元ベクトル μ 、 σ 、 z を抽出。



③ 混合ガウスモデル(8カテゴリ)でクラスタリング

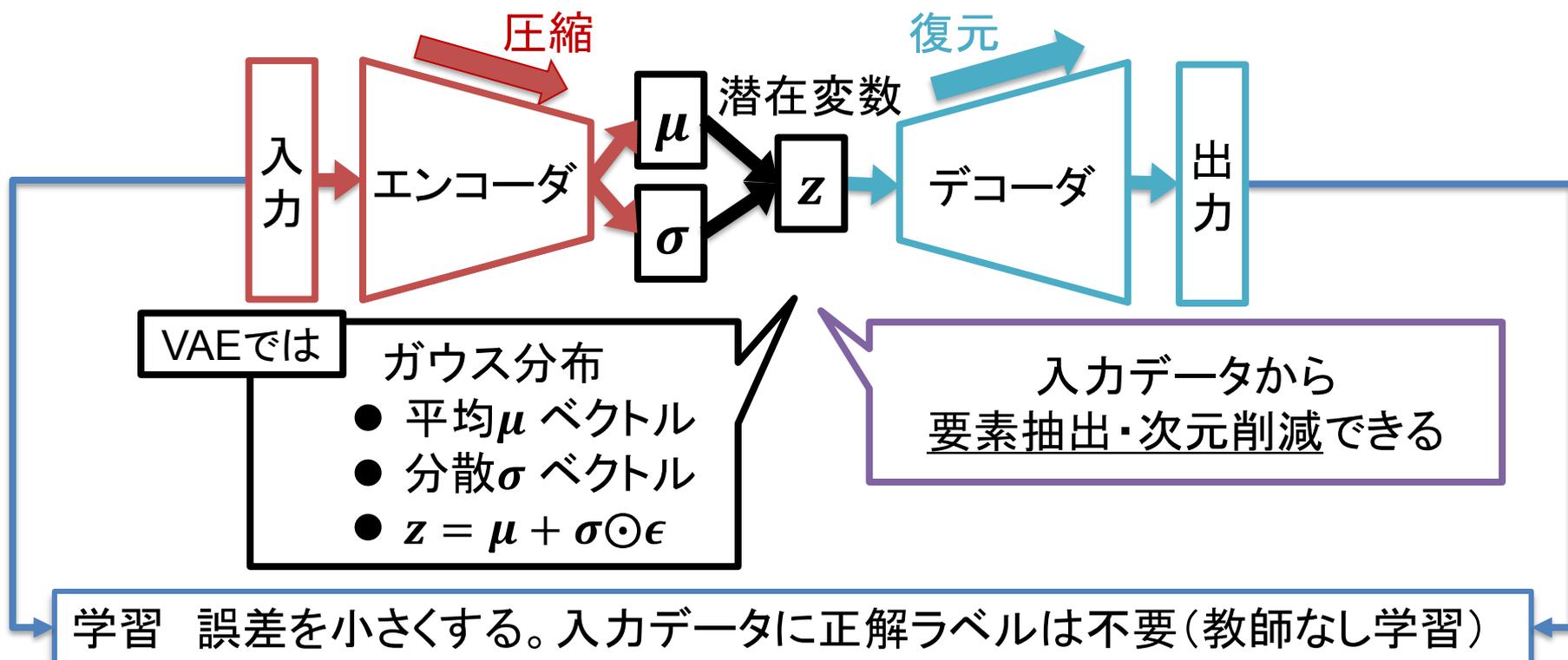
GMMはデータ分布を複数のガウス分布(各カテゴリ)の重ね合わせで近似。1個のデータ点は各カテゴリの混合で表現、負担率が最大のカテゴリで代表させる。



教師なし深層学習 変分オートエンコーダ (VAE)

観測データから未知の特徴を見つけたいので、教師なし学習を用いる。
教師なし深層学習の中で、潜在変数を構築し、次元削減のできる Variational Autoencoder (VAE; Kingma & Welling 13)を使用する。

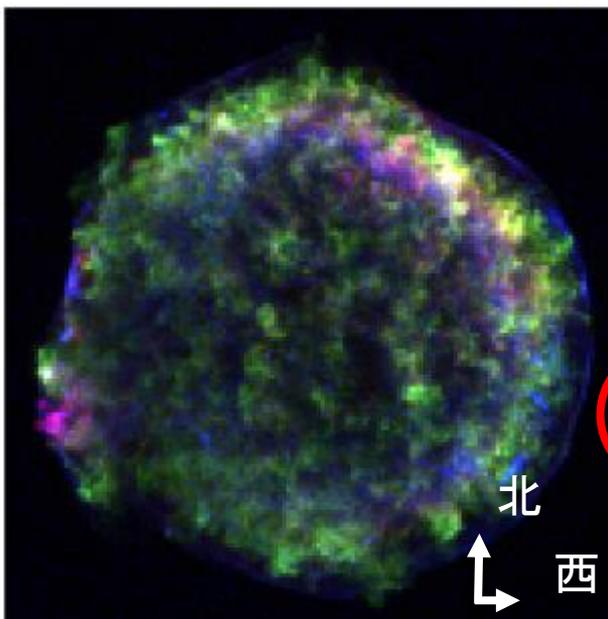
VAEは、入力をエンコーダで μ 、 σ に圧縮し、入力より低次元のガウス分布に従う z を生成する。 z からデコーダで元のデータ形状に復元する。



擬似3色画像 vs クラスタリングマップ

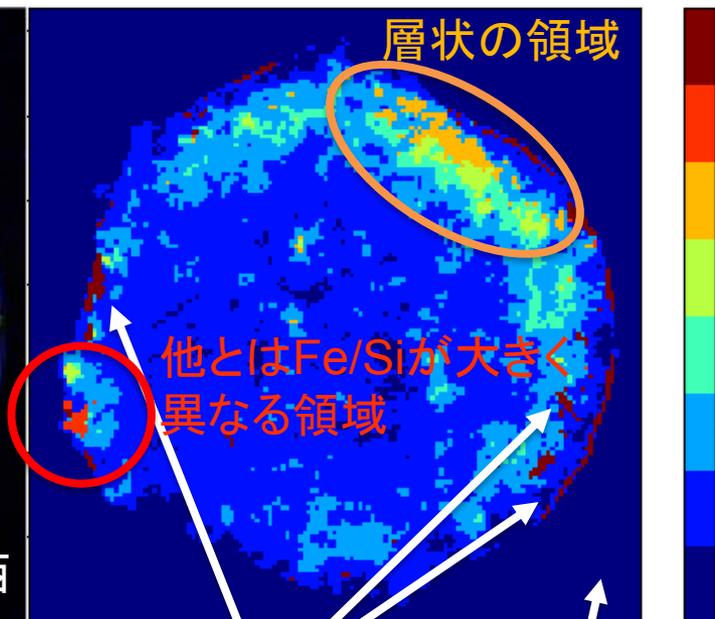
- ✓ 特徴的な構造(既知のものも含めて)を分類できている
- ✓ 空間情報を与えていないが、空間的な分類が自動でされた
- ✓ カテゴリの領域からスペクトル抽出すれば、物理的な解析が可能。

擬似3色画像



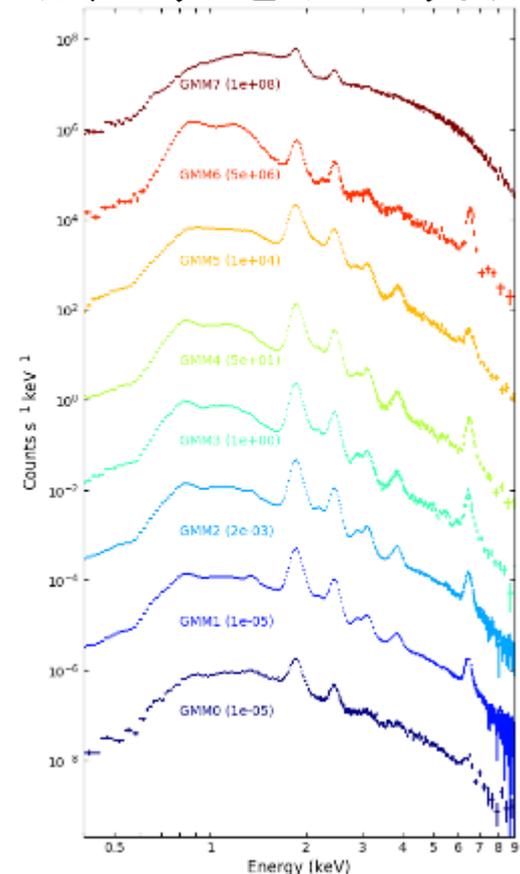
R 0.7—0.95 keV Fe L
G 1.75—1.95 keV Si He α
B 6.2—6.9 keV Fe K α

クラスタリングマップ



外縁部の衝撃波、
フィラメント構造

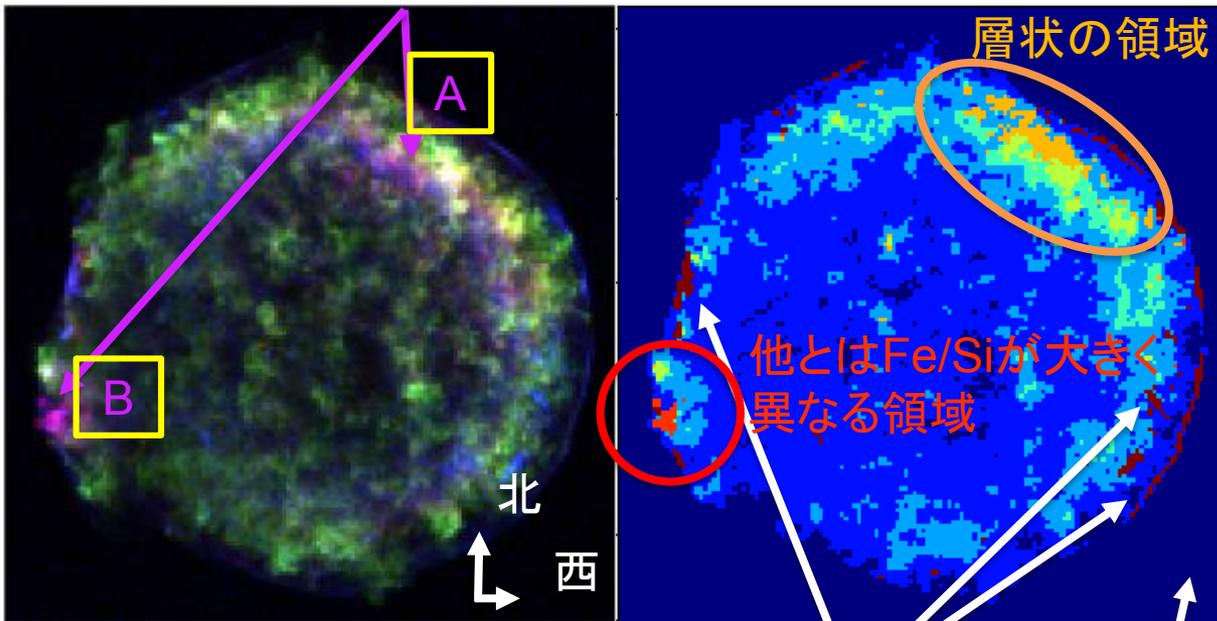
カテゴリごとのスペクトル



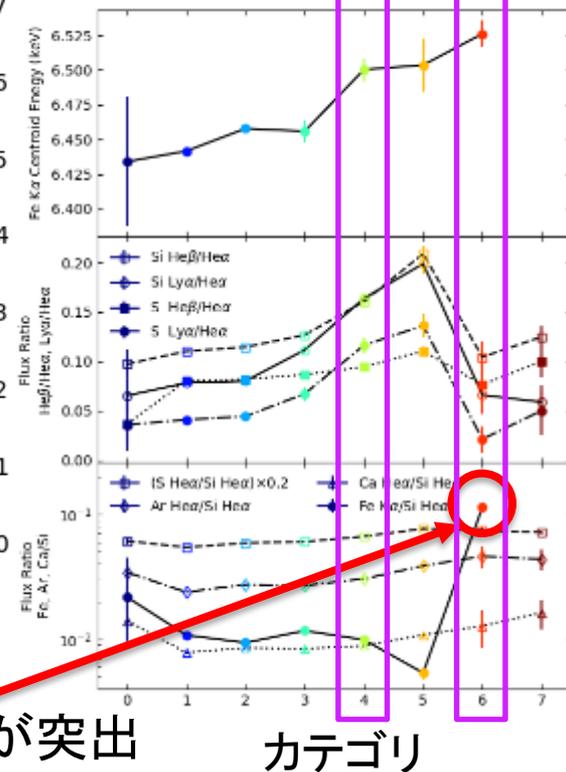
擬似3色画像 vs クラスタリングマップ

- ✓ 特徴的な構造(既知のものも含めて)を分類できている
- ✓ 空間情報を与えていないが、空間的な分類が自動でされた
- ✓ 3色画像では同じような色でも、別カテゴリに分類できている

例 同様の赤色だが、カテゴリは別々



カテゴリごとにスペクトルの特徴が異なる

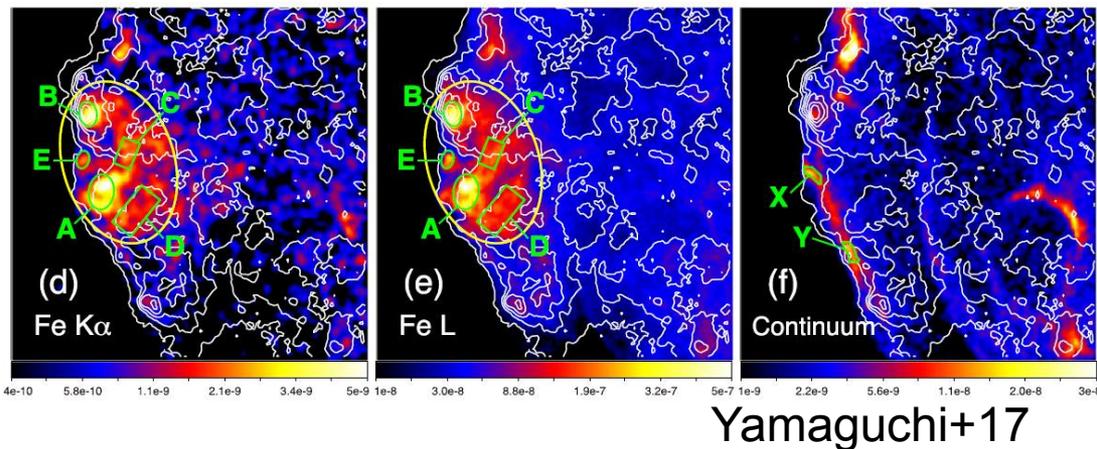


R 0.7—0.95 keV Fe L
 G 1.75—1.95 keV Si Heα
 B 6.2—6.9 keV Fe Kα

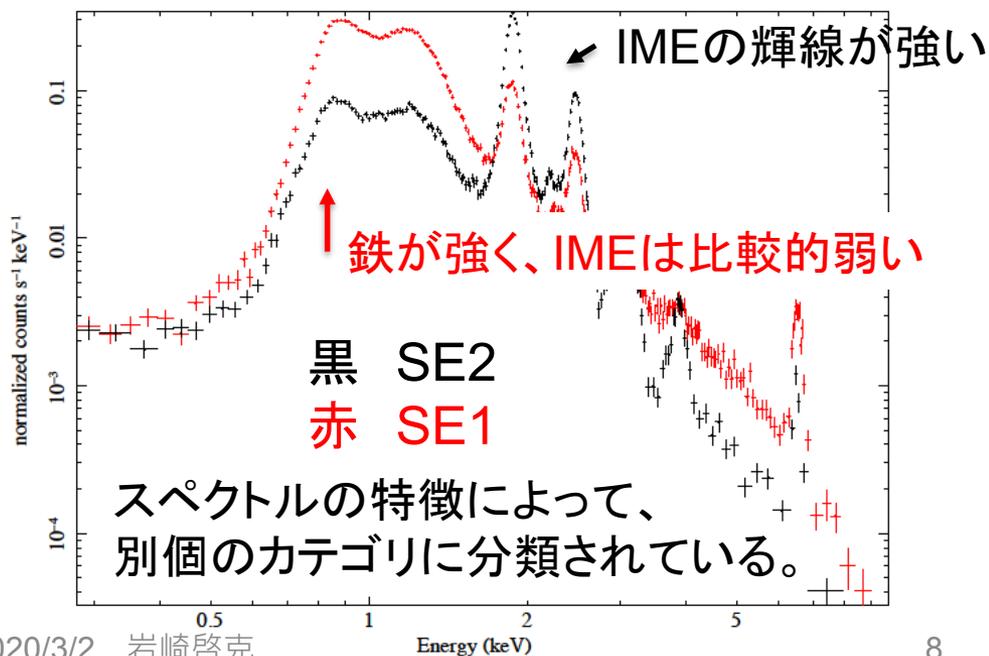
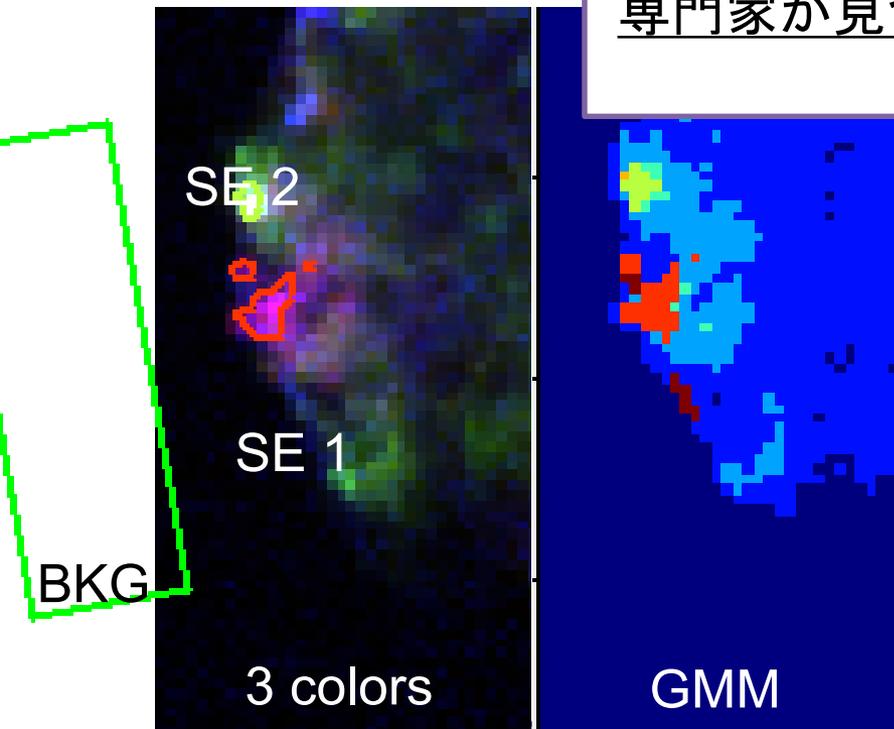
外縁部の衝撃波、
 フィラメント構造
 SNRの外側
 Fe/Si が突出

南東 Fe knot

鉄ノット (Yamaguchi+17で詳細に解析されている) の中で、論文での領域分けと同様の、空間的な分類が自動的にされた。鉄/中間質量元素 (IME) が突出して強い領域が自動的に抽出された。



専門家が見つけた物理的な解釈につながる空間構造が自動的に再発見された



Kepler' SNR への適用

VAEとGMMを組み合わせた手法をKepler's SNRの観測データにも適用した。Kepler's SNRは若いIa型SNR。濃い星周物質に囲まれ、SNRの形状が複雑。

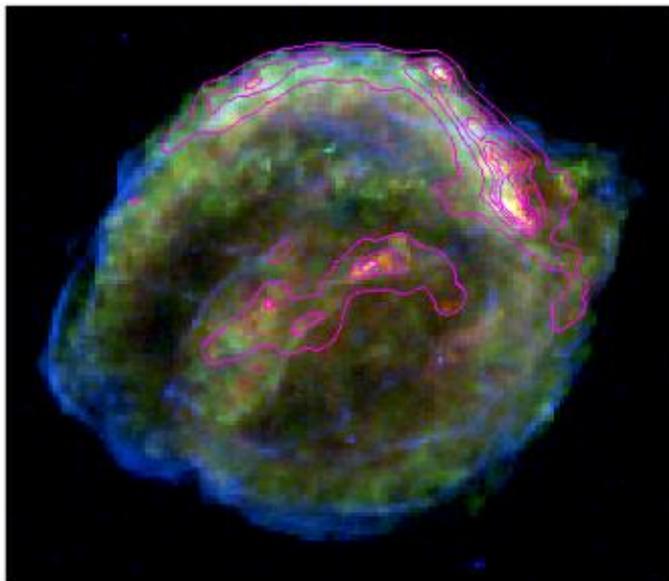
Kepler's SNR (SN 1604)

視直径 約3分角

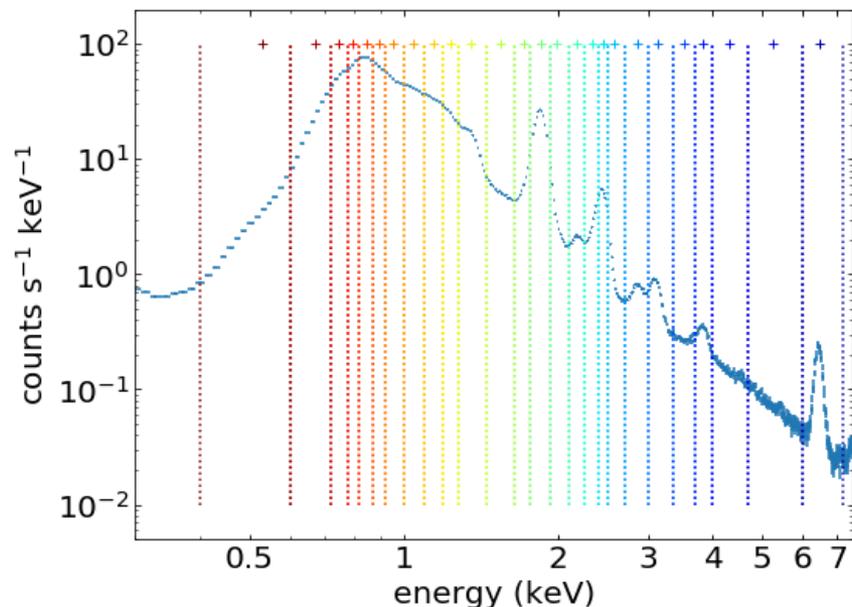
距離 約5 kpc

Chandraで975 ks観測

SNR全体のスペクトルを、各ビンに十分な光子数があり、かつ特徴的な輝線が別々のビンに分かれるように26分割。



擬似3色画像 O K; Si K; continuum;
contour, warm dust



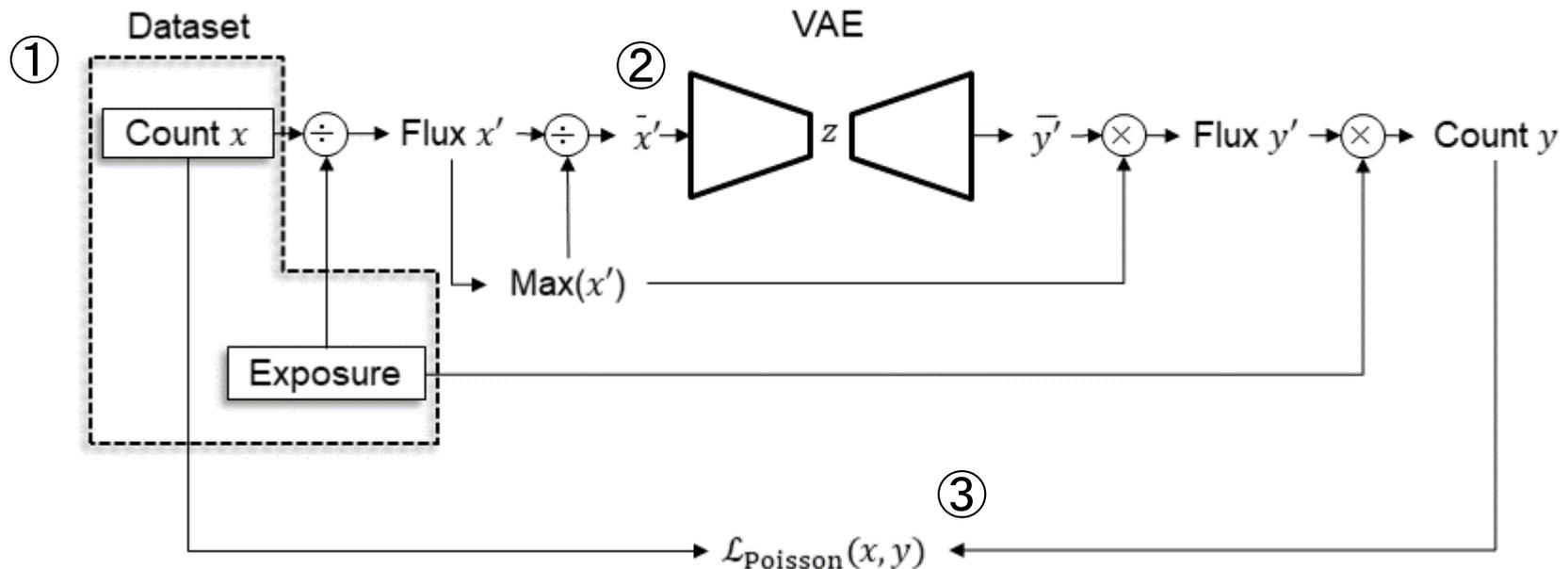
Poisson 誤差を用いたVAE

光子スペクトルはPoisson分布に従うため、他の天体にも応用できるように **Poisson likelihood** を再構成誤差に用いる。

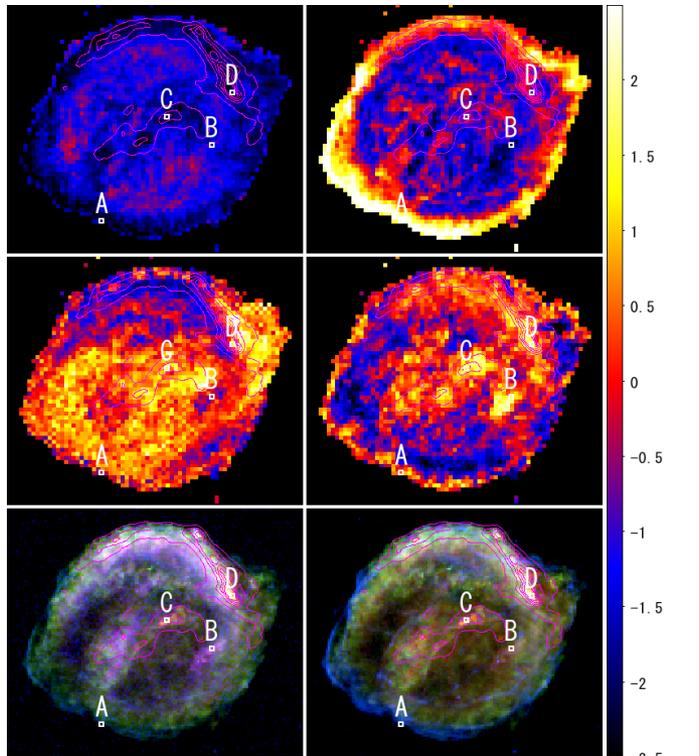
- Poisson likelihood (Ichinohe&Yamada'19) では、**カウント**を用いる。
- VAEでは**光子フラックス**のスペクトル形状を学習したい。

観測データに適用できるように新しくVAEの入出力に工夫を加えた。

- ① カウントとExposureを対にしたデータを用意。
- ② VAEへ入力時にフラックスに変換、各スペクトルの最大値で規格化。
- ③ VAEの出力をカウントに戻し、Poisson loss を計算。



VAEによるスペクトルの再現



潜在変数 z の各軸

左下 Fe L, Si K, Fe Ka

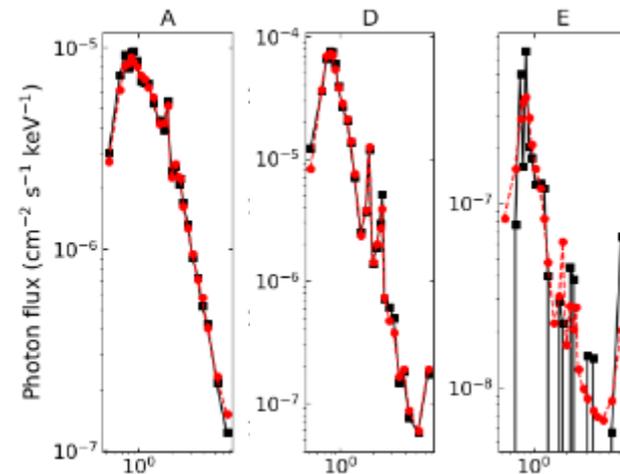
右下 O K, Si K, continuum

等高線 warm dust

VAEで再構成したスペクトルと、元のスペクトルを比較。

特徴の異なるスペクトルもそれぞれ再現されている。

低カウントによってばらつきのある大きいスペクトルはなめらかに再構成されている。

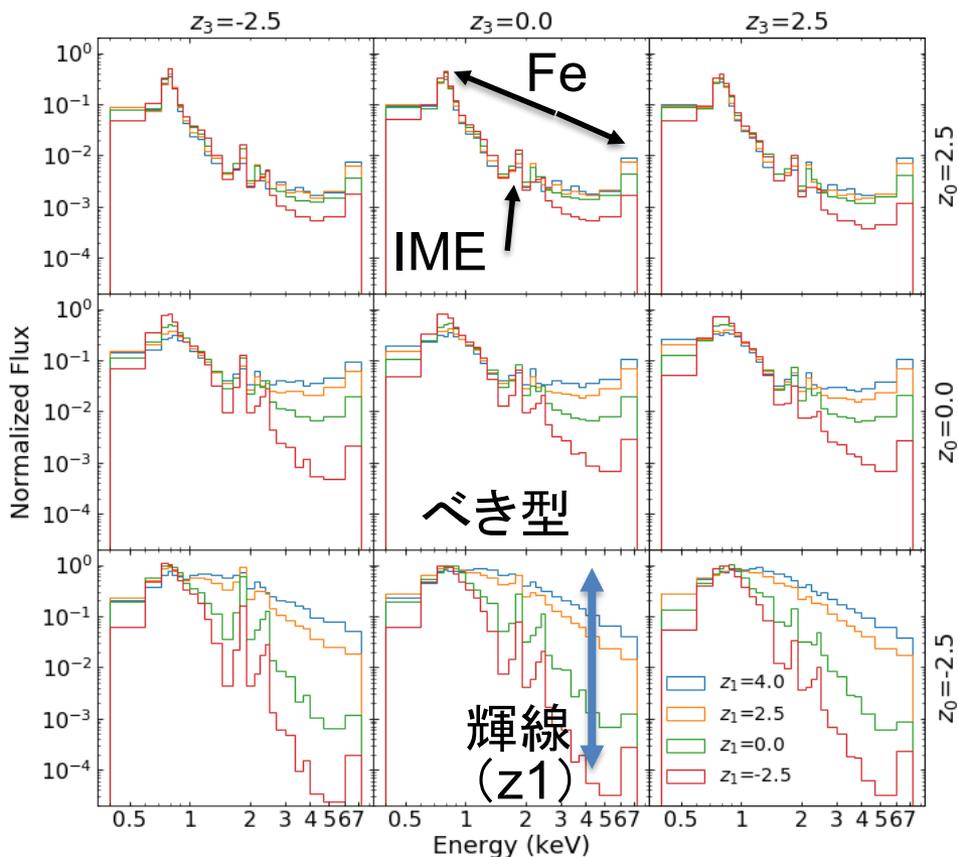


黒 元のスペクトル、赤 再構成スペクトル

デコーダの生成スペクトルを用いた潜在変数の解釈

デコーダに入力する潜在変数の値を変化させて、生成されるスペクトルの変化を見ることで、潜在変数が表すスペクトルの特徴を解釈できる。

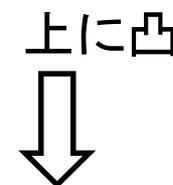
← IME輝線強い(z_3) $z_2=0.2$ 弱い →



下に凸
(z_0)

z が表すスペクトルの特徴

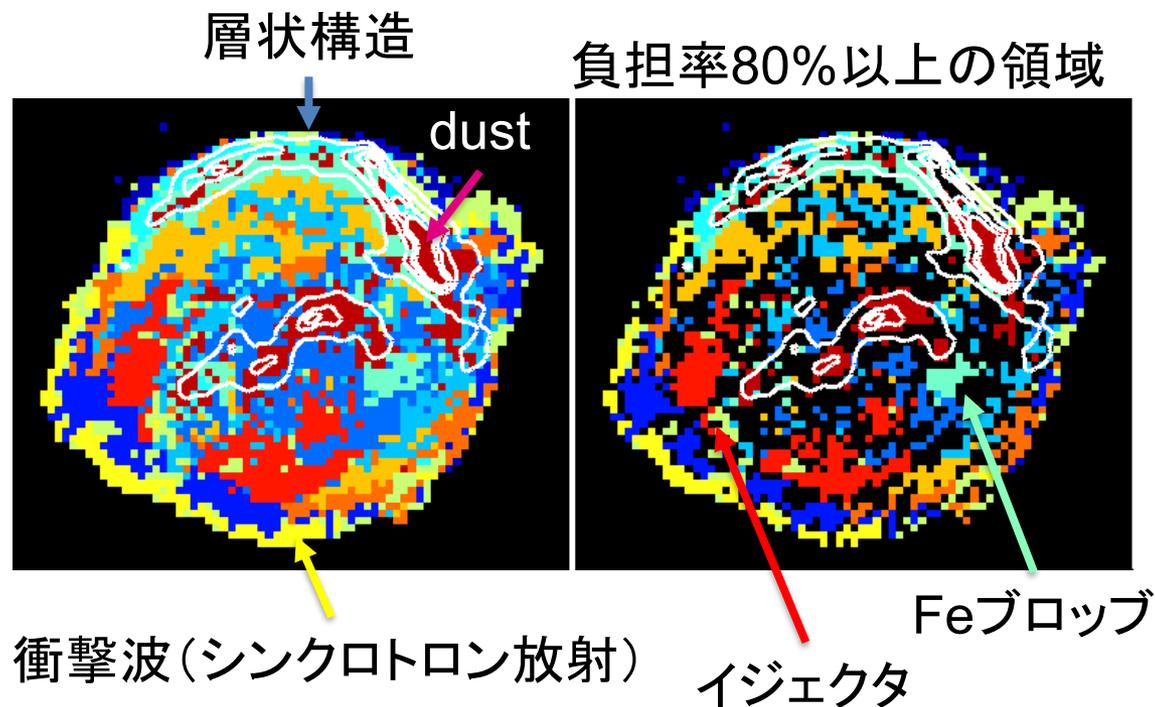
- z_0 全体の形、IME輝線と鉄輝線の関係
- z_1 連続成分と輝線成分
- z_2 軽い元素(N、O)と重い元素(S、Fe)
- z_3 IME輝線強度、鉄輝線の電離状態



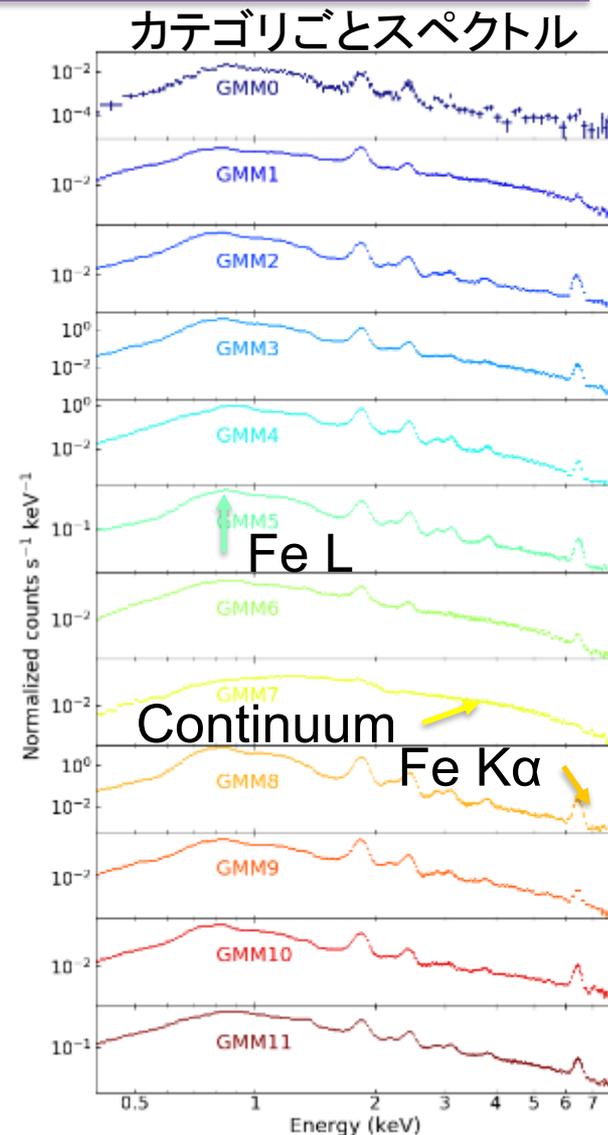
上に凸

Kepler's SNRのクラスタリングマップ

VAEの潜在変数をGMMでクラスタリングした。
特徴的な空間構造が分類されている。



濃い星周物質と相互作用している領域、
純粋なイジェクタ、衝撃波といった特徴が
自動的に抽出された。



まとめ

- VAEとGMMを組み合わせた手法を、天文データに応用した。
 - 広がった天体の空間分解した分光解析へのDNNの導入
 - 教師なし機械学習による分光観測データからの特徴抽出
- SNRのX線解析に応用し、人の手を介さずに、スペクトルの特徴に基づいて空間構造を自動的に抽出した。
 - 機械学習では空間情報を与えていないが、スペクトルの特徴にもとづいて、空間的な分類が自動的にされた
- VAE(次元削減)とクラスタリングを組み合わせた手法は観測データから特徴を抽出し分類する強力なツールとなりうる。
 - 自動的にスペクトルの特徴から空間構造を抽出
 - 教師なし学習であるため、データへのラベル付けは不要
 - 将来のX線検出器(Athena等)の高分解能データの解析にも対応可能
 - スペクトル情報のみを使うため、様々な観測に応用可能
 - 広がった天体、時間変動など。他の波長や多波長観測でも使用可