



Методы геометрической морфометрии для функциональной диагностики сердечной патологии

**Рокеах Роман Олегович, Чумарная Татьяна
Владиславовна, Соловьёва Ольга Эдуардовна**
ИИФ УрО РАН, лаборатория математической
физиологии

Мотивация



- Оценка характеристик формы сердца является важным диагностическим инструментом в кардиологии. Форма сердца влияет на сердечную функцию и наоборот.
- Эхокардиография (ЭхоКГ) - доступный метод визуализации формы сердца. Производит большие объемы информации, которая в клинической практике используется недостаточно:
 - разные проекции,
 - разные фазы в течение сердечного цикла
 - динамическое наблюдение за пациентом

Мотивация



- Геометрическая морфометрия (ГМ) - математический подход для анализа формы
- Получение параметров формы с помощью ГМ не требует большого набора данных, перебора гиперпараметров и не требует учителя, в отличие от подходов, использующих нейронные сети. Может быть эффективно в условиях ограниченных выборок и некачественной разметки.
- ГМ успешно используется в биологии, экологии, анатомии, антропологии. Однако применительно к исследованиям сердца публикации редки

Геометрическая морфометрия

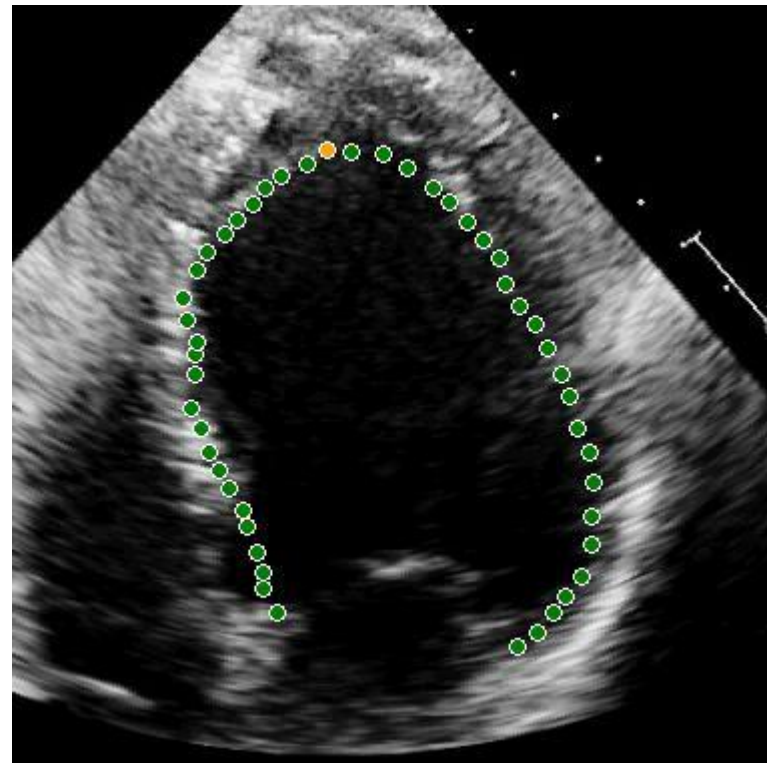


- Объекты исследования лежат в m -мерном пространстве.
- Кодировем геометрическую информацию об объектах к метками - k точками на объектах.

Конфигурация - набор меток на конкретном объекте.

Конфигурационная матрица X - это $k \times m$ матрица декартовых координат k меток в m измерениях.

- Пространство конфигураций $R(k \times m)$



Конфигурация эндокардиального контура левого желудочка сердца поверх ЭХО-изображения

Центроидный размер



- Центроидный размер конфигурации X

$$S(X) = \sqrt{\sum_{i=1}^k \|(X)_i - \bar{X}\|^2}$$

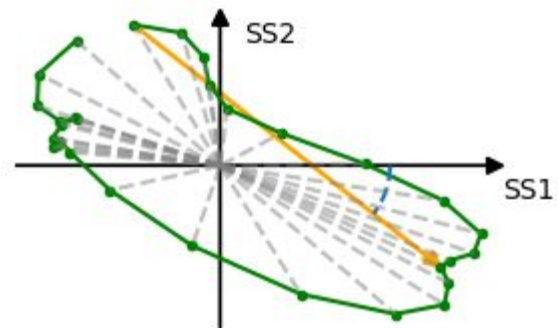


Схема признаков конфигурации. Серыми линиями обозначены отрезки от каждой точки до центра конфигурации.

Предформы



- **Предформа** Z конфигурации X - конфигурация X после переноса центра в 0 и масштабирования центроидного размера к 1 .
- **Пространство предформ**

$$S_m^k \equiv S^{(k-1)m-1}$$

это гиперсфера единичного радиуса в $(k-1)m$ -мерном пространстве.

Формы



- **Форма** $[X]$ конфигурации X - вся геометрическая информация об X , которая инвариантна при переносе, масштабировании и повороте.
 $[X] = \{Z\Gamma : \Gamma \in SO(m)\}$
- **Пространство форм** имеет размерность $q = km - m - 1 - \frac{m(m-1)}{2}$
- Для работы с координатами необходимо выбрать представителя для каждой формы, выбрав соответствующую матрицу поворота.

Расстояния между формами

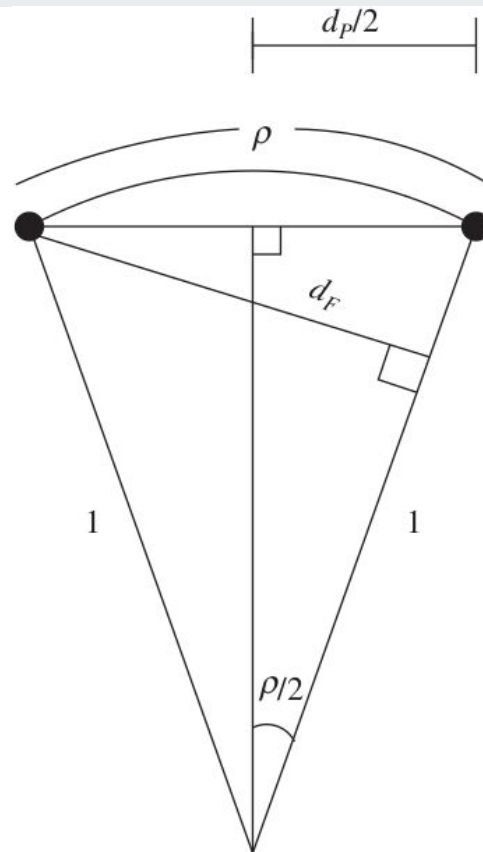
- Частичное прокрустово расстояние

$$d_P(X_1, X_2) = \inf_{\Gamma \in SO(m)} \|Z_2 - Z_1 \Gamma\|$$

- Полное прокрустово расстояние

$$d_F(X_1, X_2) = \inf_{\Gamma \in SO(m), \beta \in \mathbb{R}^+} \|Z_2 - \beta Z_1 \Gamma\|$$

- Риманово расстояние $\rho(X_1, X_2)$
наименьшая дуга большого круга
между Z_1 и Z_2 на сфере предформ
- $\|\cdot\|$ - норма Фробениуса
- Для расстояний и поворотов
выведены закрытые формулы



Сечение сферы предформ с различными расстояниями между двумя точками.

Выбор представителей формы



- Обычно выбирается средняя форма для набора данных. Затем для всех предформ выполняется оптимальный поворот к средней.

- **Средняя выборочная форма**

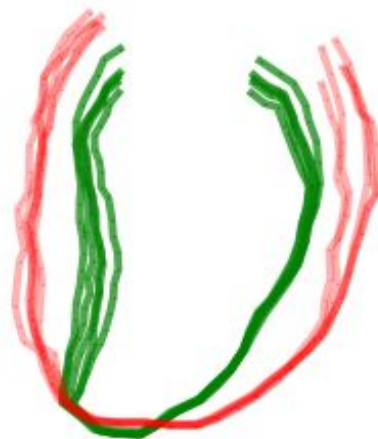
$$\hat{\mu}_M = \arg \inf_{\mu} \frac{1}{n} \sum_{i=1}^n d^2(X_i, \mu)$$

- Для нахождения средней формы используется алгоритм **обобщенного прокрустового анализа** (GPA - generalized procrustes analysis).

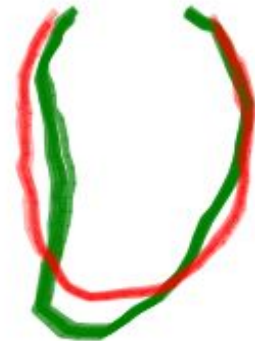
Обобщенный прокрустов анализ (GPA)

1. Даны матрицы предформ Z_i .
2. Инициализировать среднее. Например, $\mu_1 = Z_0$.
3. Для каждого i выполнить оптимальный поворот и масштабирование между Z_i и μ_1 , сохранить в Z_i .
4. $\mu_0 = \mu_1$; $\mu_1 = \text{mean}(Z_i)$.
5. Повторять шаги 3,4 до выполнения остановочного критерия. Например, $\|\mu_1 - \mu_0\| < \text{atol}$.

До трансформаций



После GPA



Контурсы эндокарда левого желудочка до и после применения GPA. Слева исходные конфигурации, справа представители формы вокруг среднего

Пример 1.

Выборка

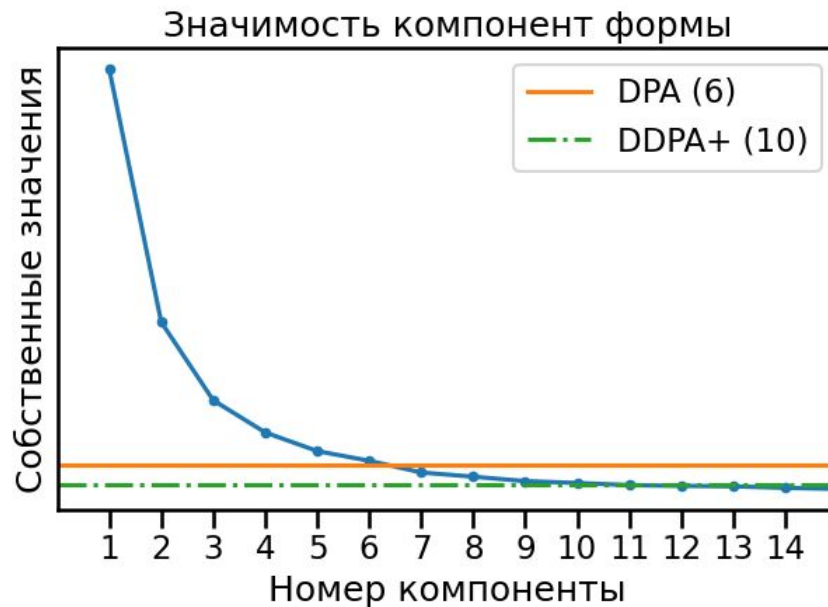
- Видеозаписи эхокардиографического исследования ЛЖ в 4-камерной позиции в течение сердечного цикла. (**n=120**)
- **Здоровые волонтеры n=34**
- **Пациенты с диагнозом хроническая сердечная недостаточность (ХСН, CHF) n=34**
- **Пациенты с диагнозом ишемическая болезнь сердца и сохраненной фракцией выброса (ИБС, IHD) n=52**
- Задача: Возможно ли разделение этих 3 групп по параметрам формы?

Базовая предобработка

- На видео полуавтоматически отслеживались **49 точек** контура эндокарда ЛЖ.
- Траектории точек контура интерполированы во времени для выравнивания моментов времени между разными записями (**85 временных точек**). Время **0** и **84** соответствует **конечной диастоле**, **29** - **конечной систоле**.
- Траектории аппроксимированы рядами Фурье для обеспечения непрерывности.
- Итоговый тензор имеет размер **(120, 85, 49, 2)**.

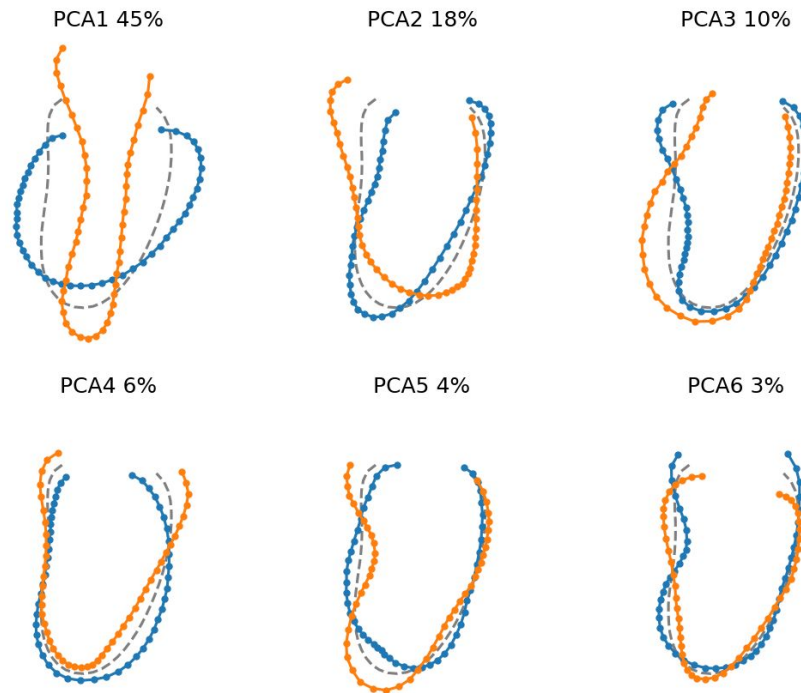
Получение параметров формы

- Выборка разделена на обучающую и тестовую части в соотношении **2 : 1** ($n_{train}=80$, $n_{test}=40$) с сохранением долей классов
- По обучающей выборке построено пространство формы ЛЖ, выполнена проекция на касательное пространство
- К точкам в касательном пространстве применен **метод главных компонент**
- Все преобразования применяются к 80×85 матрицам размера (49,2).
- Проведен тест на **значимость собственных значений** матрицы ковариаций [*]
- Оставлены **6 значимых** компонент формы



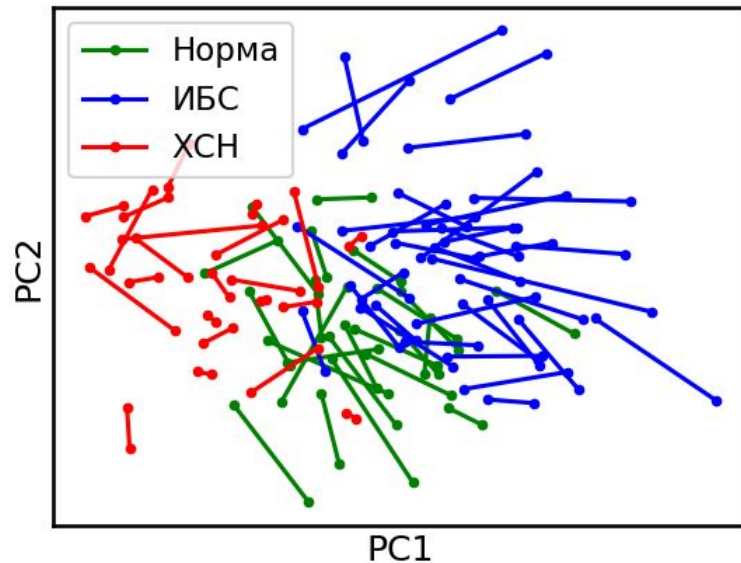
Получение параметров формы

- Визуализируем формы компоненты формы в исходных координатах. Синим представлено минимальное значение соответствующей компоненты формы, оранжевым - максимальное. (гиперболизировано на 1.5).
- Компоненты формы связаны с разными частями стенки ЛЖ и разными изменениями формы.

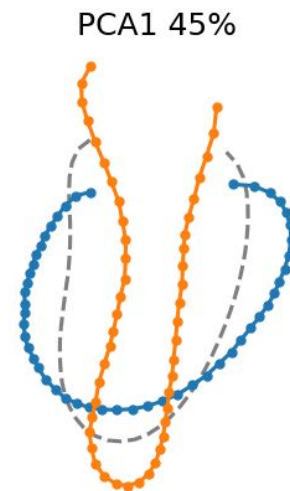
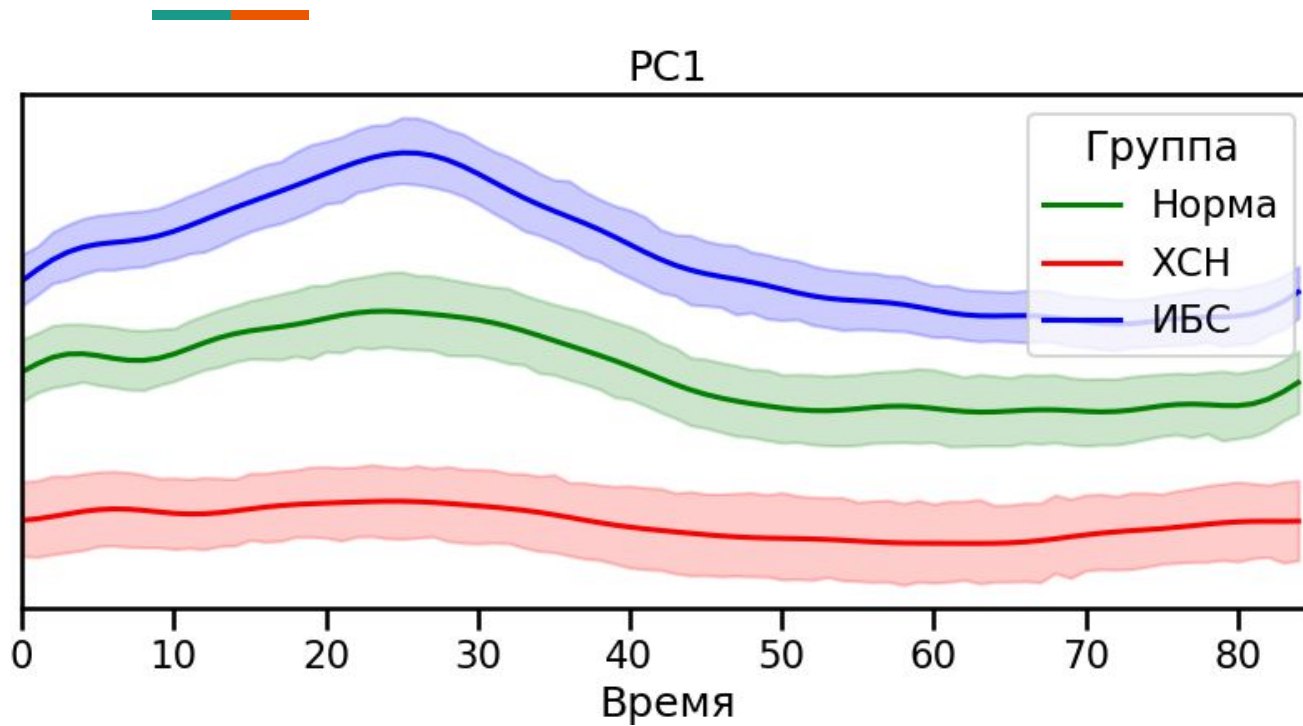


Выбор параметров для классификации

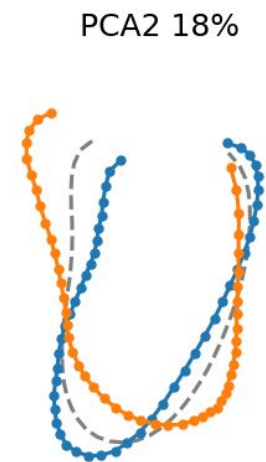
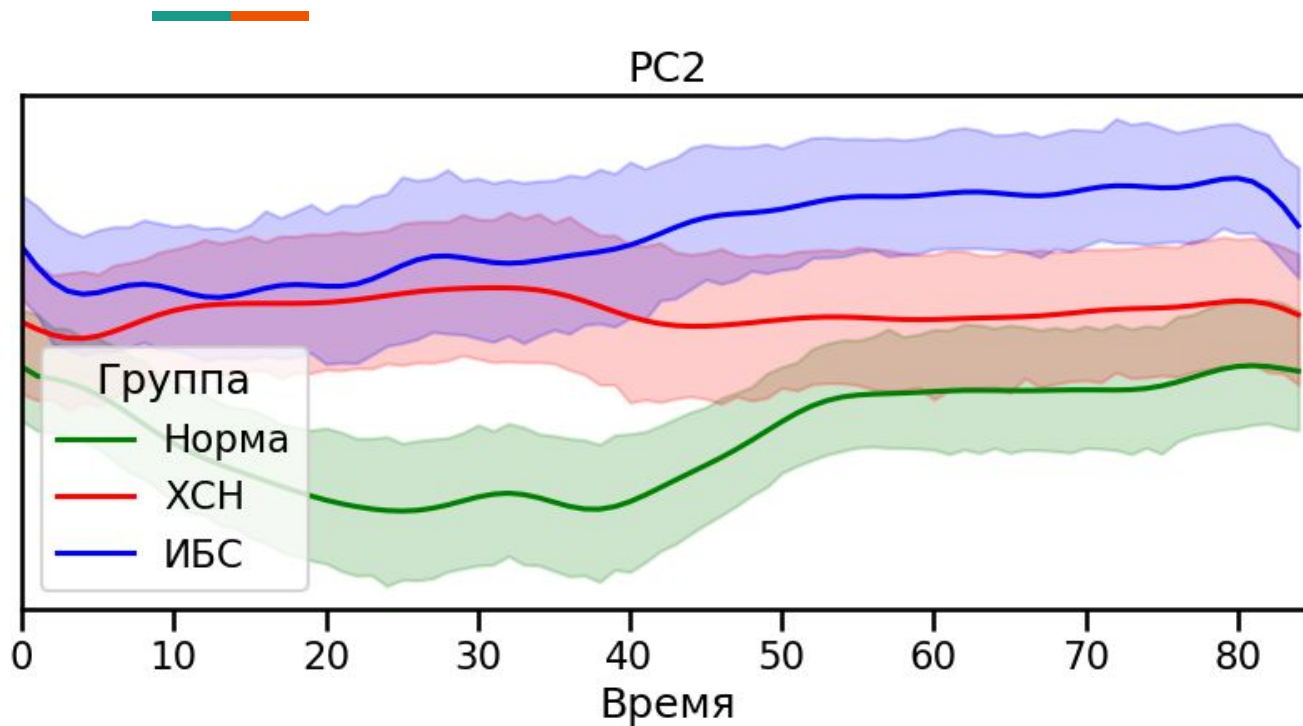
- Визуализируем формы ЛЖ в проекции на первые две компоненты пространства (временные точки КД и КС)
- Наблюдаем отчетливое разделение групп в построенной проекции
- Проведем групповые сравнения между значениями компонент формы в каждый момент времени, чтобы выбрать признаки для построения классификатора



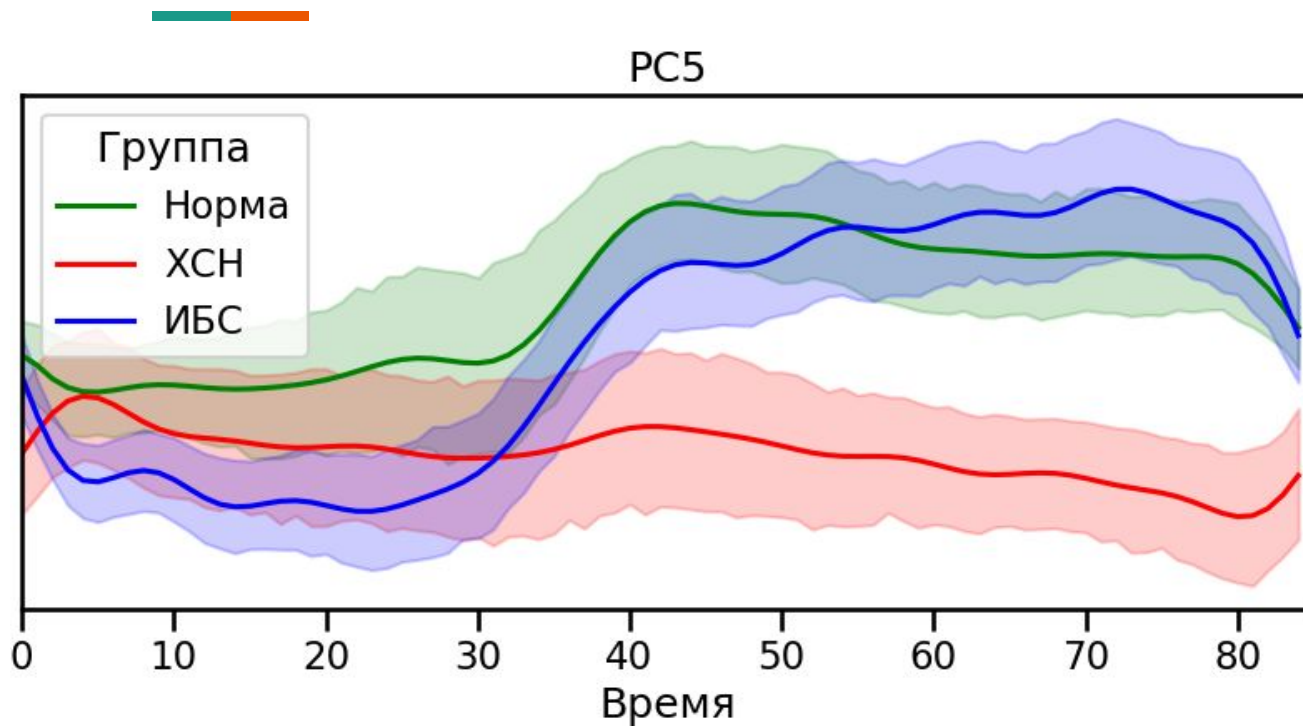
Выбор параметров для классификации



Выбор параметров для классификации



Выбор параметров для классификации



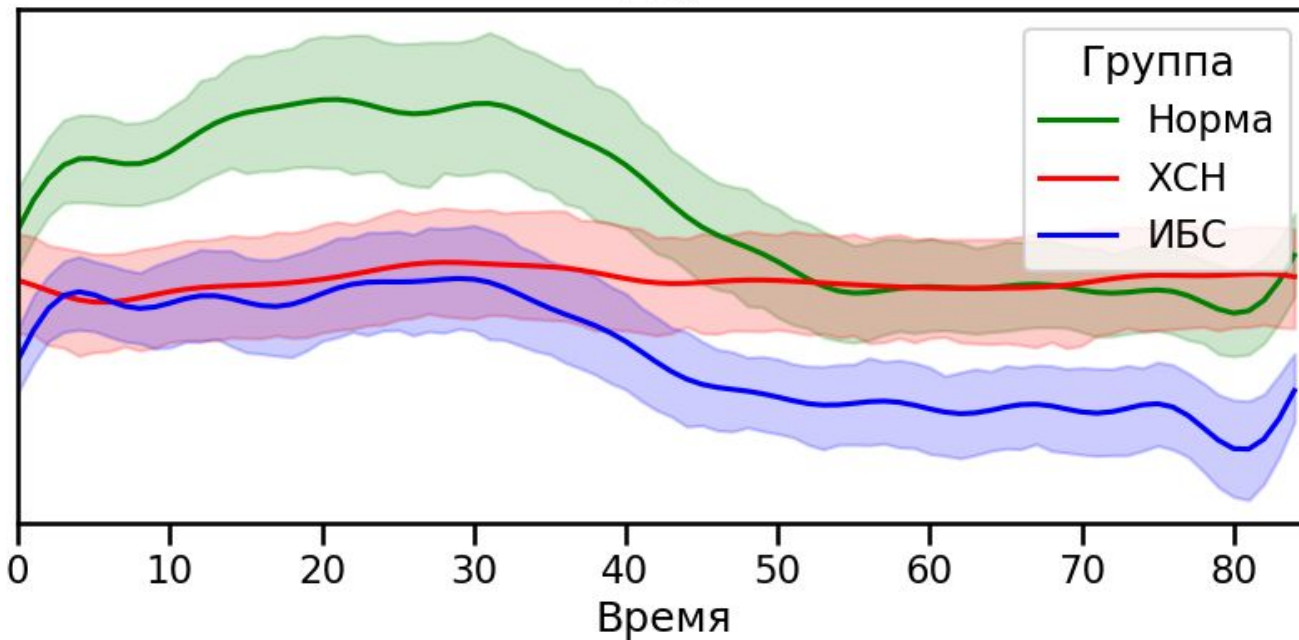
PCA5 4%



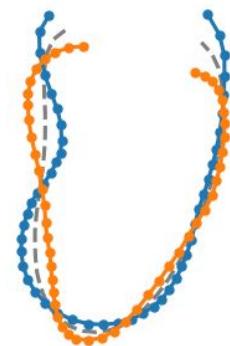
Выбор параметров для классификации



PC6

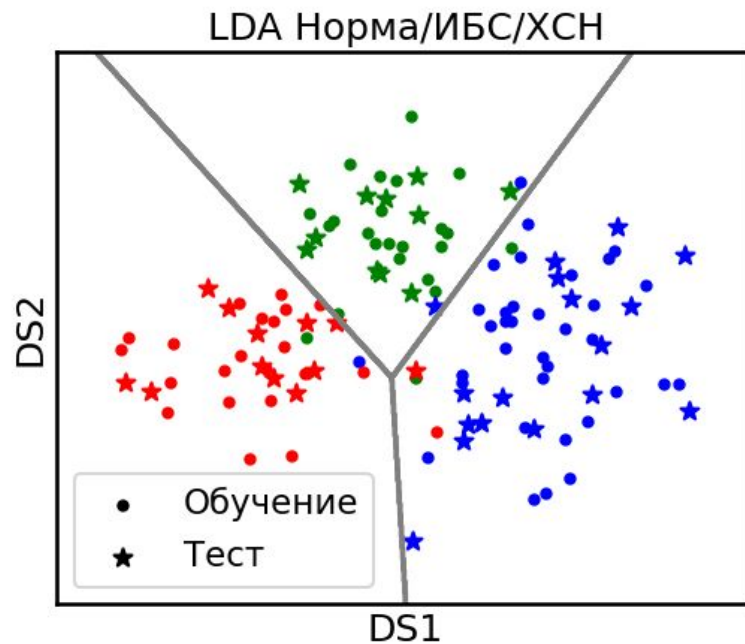


PCA6 3%



Классификация

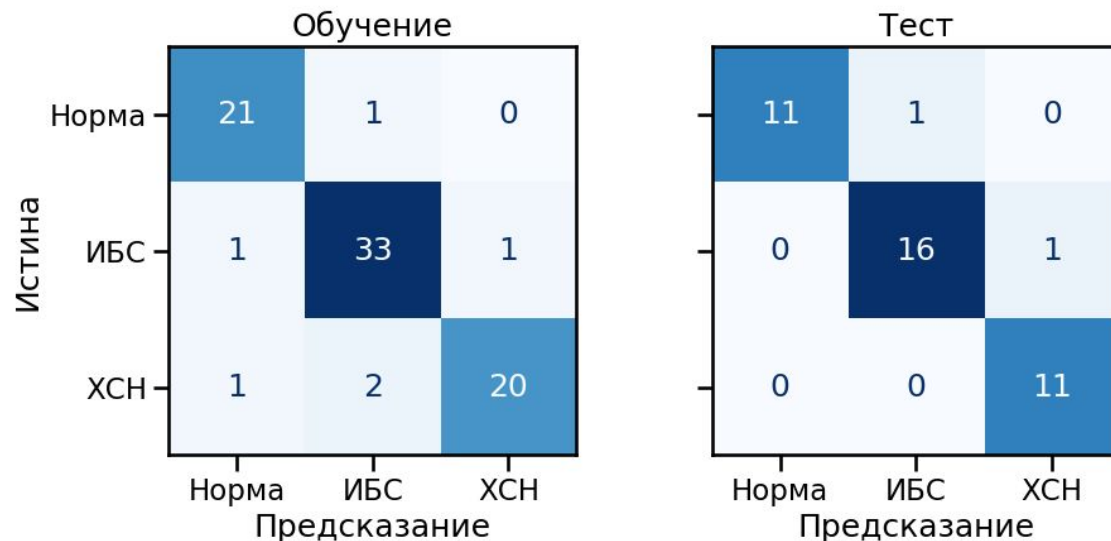
- Для классификации выбраны 6 параметров: **PC1_F29** в конечную систолу (разделяются все), **PC2_F29** в конечную систолу (норма), **PC5_F70** в позднюю диастолу (XCH), **PC6_F29** в конечную систолу (норма), **PC6_F70** в позднюю диастолу (ИБС).
- Модель классификации - линейный дискриминантный анализ
- Тестовые данные спроектированы в пространство, построенное на обучающей выборке



Классификация

- Precision (weighted avg): 0.93, 0.95
- Recall (weighted avg): 0.93, 0.95
- 74 правильных ответов из 80 в обучающей выборке
- 38 правильных ответов из 40 в тестовой выборке

LDA матрицы ошибок



Промежуточный вывод



- Методы геометрической морфометрии успешно применены для отделения ИБС с сохраненной фракции выброса от нормы
- Опробуем их для сложной задачи прогнозирования улучшения состояния после установки СРТ-кардиостимулятора

Пример 2.

Выборка

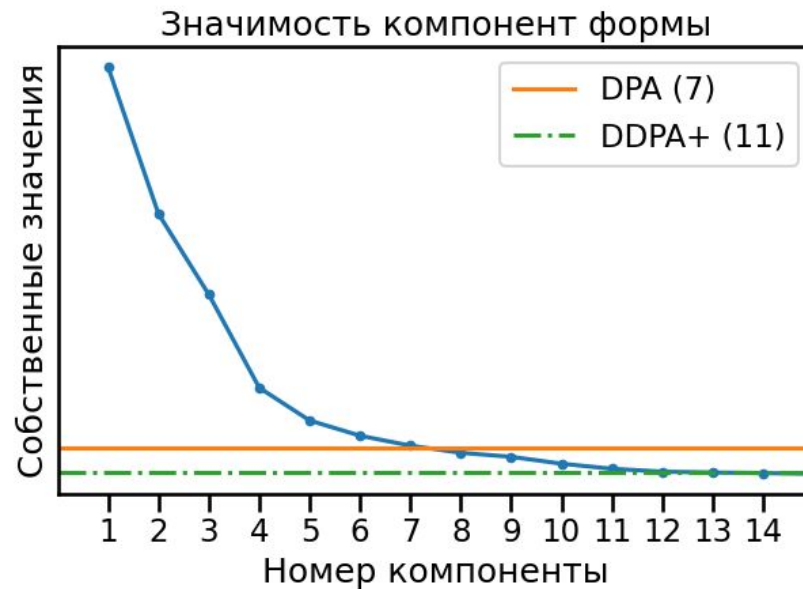
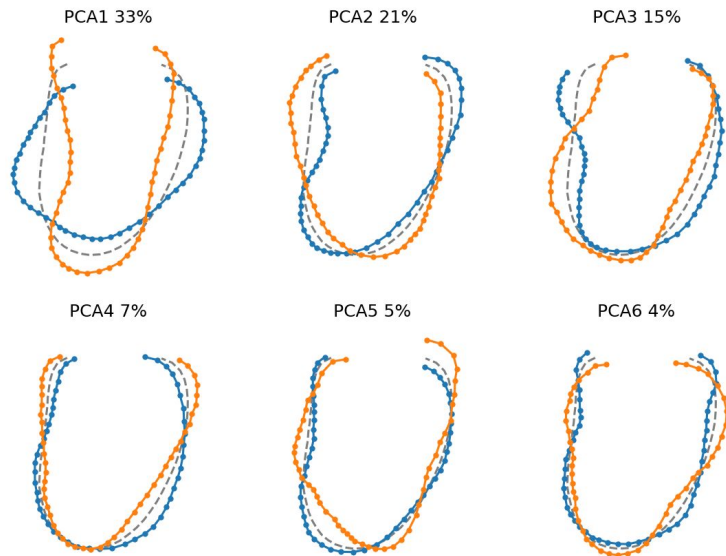
- Видео ЭхоКГ исследования ЛЖ в 4-камерной позиции в течение сердечного цикла. (**n=26**)
- Пациенты с диагнозом **хроническая сердечная недостаточность (ХСН, CHF)**, направленные на операцию для установки СРТ-кардиостимулятора.
- **Ответчики (n=13)** - пациенты, у которых улучшились ЭхоКГ показатели объемов сердца через 6 месяцев. $dКСО > 15\%$, $dФВ > 5\%$
- **Неответчики (n=13)** - пациенты, состояние которых не улучшилось.
- Задача: Возможно ли предсказать улучшение по параметрам формы, полученным из дооперационных видео?

Базовая предобработка

- На видео полуавтоматически отслеживались **49 точек** контура эндокарда ЛЖ.
- Траектории точек контура интерполированы во времени для выравнивания моментов времени между разными записями (**85 временных точек**). Время **0** и **84** соответствует **конечной диастоле**, **29** - **конечной систоле**.
- Траектории аппроксимированы рядами Фурье для обеспечения непрерывности.
- Итоговый тензор имеет размер **(26, 85, 49, 2)**.

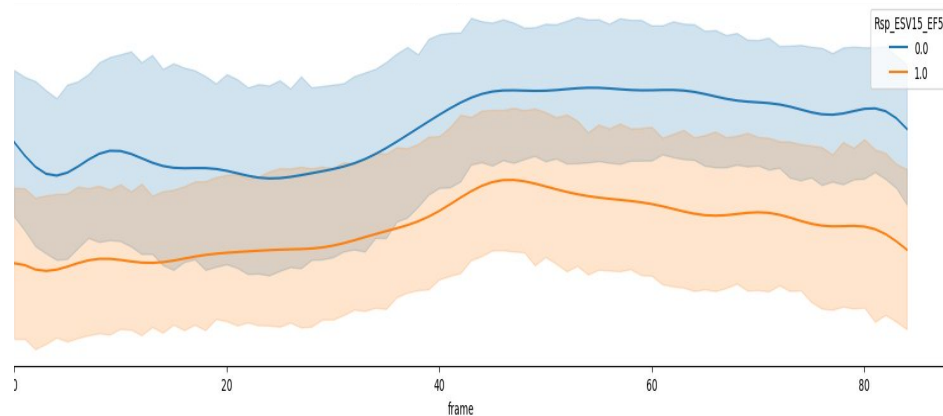
R/B/M Math et al. "Statistical Shape Analysis Reveals Disease-Specific Features of Left Ventricular Response to an Inotropic Response to Cardiac Resynchronization Therapy. Circulation. 2014;129:1111-1119.

Получение параметров формы



Получение параметров формы

- При групповых сравнениях найдено только различие в четвертой компоненте
- Что делать, когда группы не кластеризуются по полученным параметрам?
- Отличия могут быть более выражены при сравнении относительных изменений формы во времени, а не при сравнении положений траекторий относительно друг друга



Изменение PC4 в течение сердечного цикла

Параллельный перенос



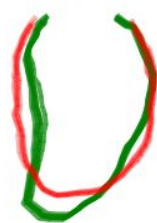
- Перенос траекторий вокруг различных полюсов пространства формы в один полюс
- Устраняет индивидуальные различия между траекториями и выделяет внутренние различия
- Аналитическое выведение точной формулы возможно, но затруднительно
- Можно воспользоваться предположениями о достаточной близости форм друг к другу и использовать линейное преобразование - **линейный сдвиг (linear shift, LS) [*]**

Линейный сдвиг

1. С помощью GPA найти глобальный полюс μ_G для всех форм из всех траекторий.
2. Внутри каждой траектории выбрать локальный полюс μ_L . Это может быть среднее или определенный момент времени.
3. Каждый μ_L оптимально повернуть к μ_G .
4. Для каждой траектории каждую конфигурацию из траектории повернуть к соответствующему μ_L .
5. Вычесть из каждой конфигурации соответствующий μ_L и прибавить μ_G .

Линейный сдвиг выполнен и далее возможно обычное применение GPA+PCA.

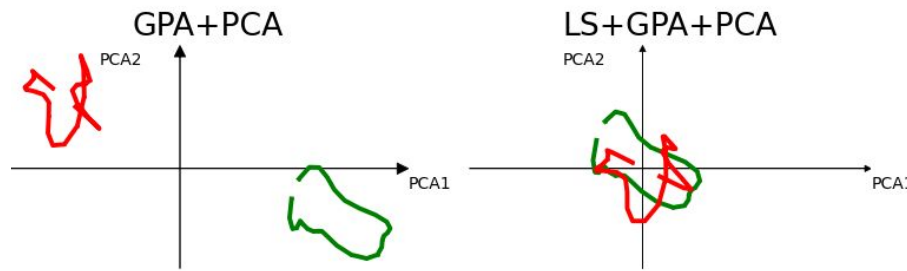
После GPA



После LS+GPA

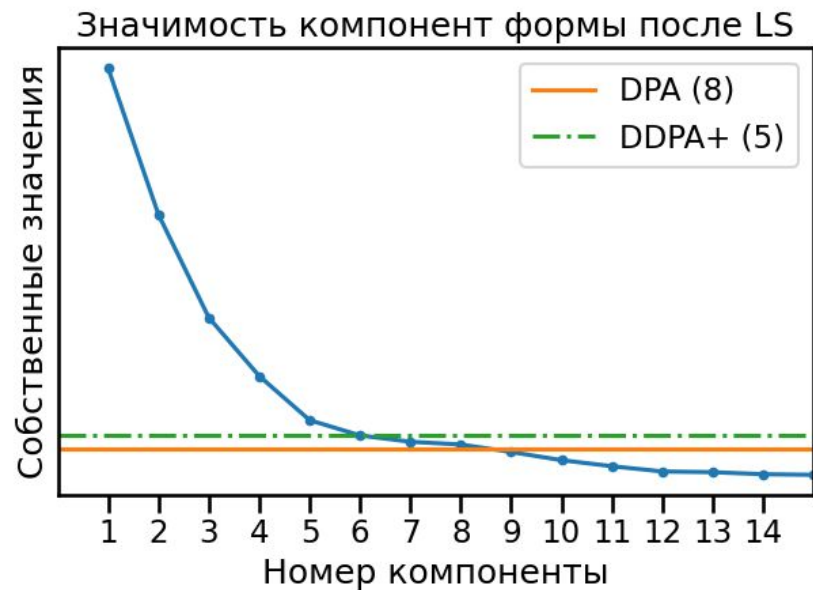
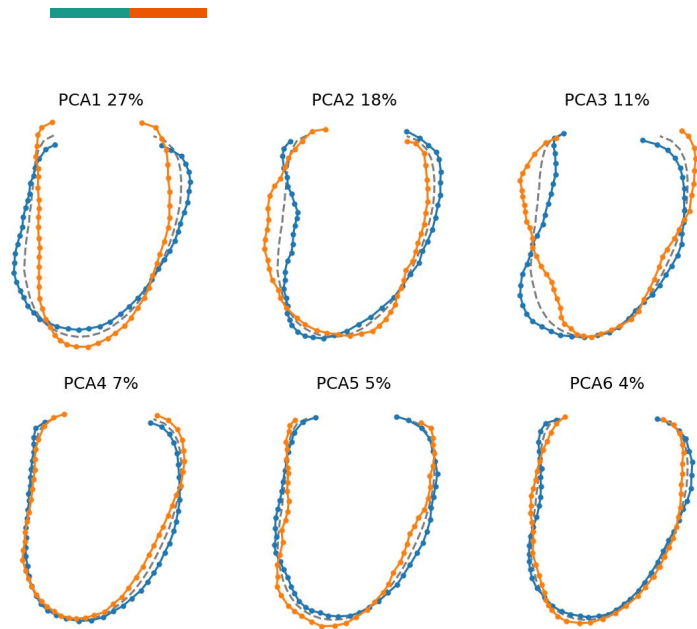


Контурные двух сердечных циклов в исходном пр-ве без и с применением LS



Траектории двух сердечных циклов в касательном пр-ве без и с применением LS

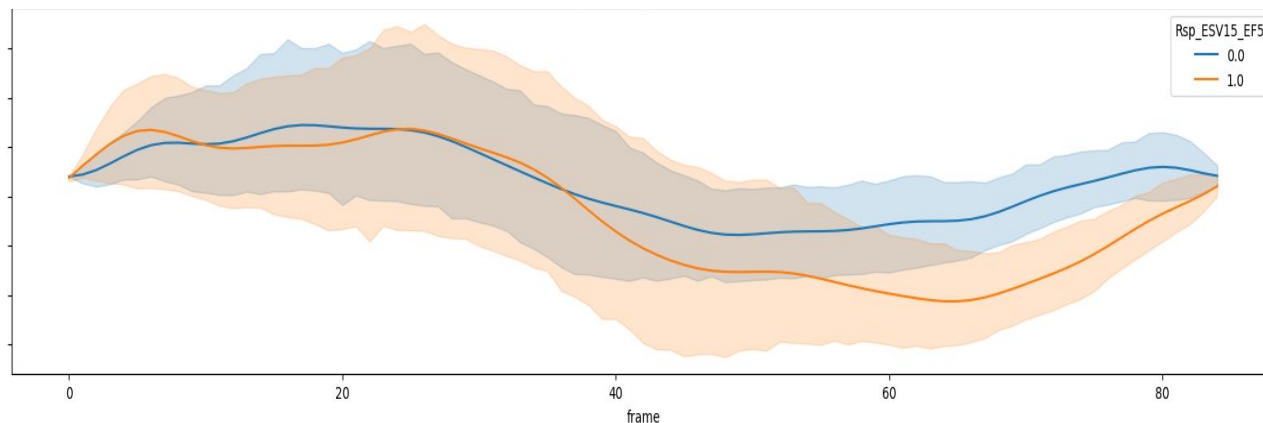
Получение параметров формы после LS



Получение параметров формы после LS



- Первая компонента значительно отличается между группами в диастолической фазе сердечного цикла

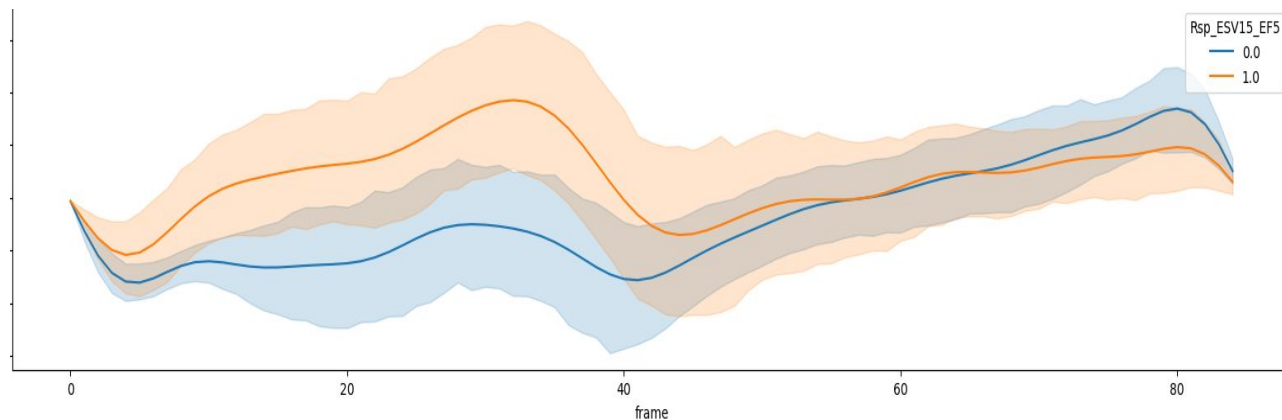


Изменение LSPC1 в течение сердечного цикла

Получение параметров формы после LS



- Вторая компонента отличается между группами в систолической фазе сердечного цикла



Изменение LSPC2 в течение сердечного цикла

Классификация



- Для классификации выбраны 3 параметра: **PC4_F0**, **LSPC1_F70** в позднюю диастолу, **LSPC2_F29** в конечную систолу
- Модель классификации - логистическая регрессия
- Значимые коэффициенты только при двух параметрах формы после линейного сдвига
- Accuracy 0.92; Precision - 0.93; Recall - 0.93; F1 - 0.93
- Добавить количество ошибок и точность при Leave-One-Out кросс

Классификатор из литературы (с таким же критерием ответа на CPT) [Feeny et al, 2019] по клиническим признакам не продемонстрировал удовлетворительной точности: Accuracy 0.54; Precision - 0.55; Recall - 0.79; F1 - 0.65

Промежуточный вывод

- Методы ГМ позволили построить классификатор прогноза ответа на СРТ более точный, чем классификаторы на клинических признаках.
- Однако размер выборки слишком ограничен, чтобы судить об обобщающей способности классификатора.

Заключение

- Мы продемонстрировали применимость методов геометрической морфометрии на задачах классификации сердечной патологии и прогноза успешности терапии.
- Возможности ГМ не ограничены извлечением признаков для задач машинного обучения. В настоящее время мы разрабатываем протокол сравнительного анализа формы сердец новорожденных с нормальным весом и в условиях недоношенности. Цель: программно реализовать обоснованную схему статистического исследования, которая должна обнаруживать значимые различия формы на уровне сегмента стенки сердца и/или фазы сердечного цикла.